Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **487,** 198–227 (2019)
Advance Access publication 2019 May 4

doi:10.1093/mnras/stz1216

The evolution of SMBH spin and AGN luminosities for $z < 6$ within a semi-analytic model of galaxy formation

Andrew J. Griffin [1]★ Cedric G. Lacey [1] Violeta Gonzalez-Perez [1,2] Claudia del P. Lagos [3,4] Carlton M. Baugh[1] and Nikos Fanidakis[5,6]

[1]*Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK*
[2]*Institute of Cosmology and Gravitation, University of Portsmouth, Burnaby Road, Portsmouth PO1 3FX, UK*
[3]*International Centre for Radio Astronomy Research, University of Western Australia, 35 Stirling Highway, Crawley WA 6009, Australia*
[4]*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Canberra, ACT 2611, Australia*
[5]*Max-Planck-Institute for Astronomy, Königstugl 17, D-69117 Heidelberg, Germany*
[6]*BASF, Carl-Bosch Strasse 38, D-67056 Ludwigshafen, Germany*

Accepted 2019 April 30. Received 2019 April 26; in original form 2018 June 21

ABSTRACT

Understanding how active galactic nuclei (AGNs) evolve through cosmic time allows us to probe the physical processes that control their evolution. We use an updated model for the evolution of masses and spins of supermassive black holes (SMBHs), coupled to the latest version of the semi-analytical model of galaxy formation GALFORM, using the *Planck* cosmology and a high-resolution Millennium style dark matter simulation to make predictions for AGN and SMBH properties for $0 < z < 6$. We compare the model to the observed black hole mass function and the SMBH versus galaxy bulge mass relation at $z = 0$, and compare the predicted bolometric, hard X-ray, soft X-ray, and optical AGN luminosity functions to observations at $z < 6$, and find that the model is in good agreement with the observations. The model predicts that at $z < 2$ and $L_{\rm bol} < 10^{43}$ erg s^{-1}, the AGN luminosity function is dominated by objects accreting in an advection-dominated accretion flow disc state, while at higher redshifts and higher luminosities the dominant contribution is from objects accreting via a thin disc or at super-Eddington rates. The model also predicts that the AGN luminosity function at $z < 3$ and $L_{\rm bol} < 10^{44}$ erg s^{-1} is dominated by the contribution from AGNs fuelled by quiescent hot halo accretion, while at higher luminosities and higher redshifts, the AGN luminosity function is dominated by the contribution from AGNs fuelled by starbursts triggered by disc instabilities. We employ this model to predict the evolution of SMBH masses, Eddington ratios, and spins, finding that the median SMBH spin evolves very little for $0 < z < 6$.

Key words: galaxies: active – galaxies: high-redshift – quasars: general.

1 INTRODUCTION

Ever since quasars were first identified to be cosmological sources (Schmidt 1968), a key aim has been to understand their evolution through cosmological time. Early studies showed that the number density of quasars shows strong evolution, with more luminous quasars present at $z \approx 2$ than at $z \approx 0$, leading to the suggestion that quasars evolve by 'pure luminosity evolution'. In this scenario, quasars are long lived and fade through cosmic time, leading to an evolution in the luminosity function of only the characteristic luminosity (e.g. Boyle et al. 1990). However, more recent optical surveys, which can probe both the faint and bright ends of the luminosity function, have shown not only that the slope of the luminosity function evolves (e.g. Richards et al. 2006; Croom et al. 2009) but also that the number density decreases at high redshift (e.g. Fan et al. 2001; Jiang et al. 2016). Surveys at X-ray wavelengths show an evolution in the shape of the luminosity function (e.g. Ueda et al. 2014) as well as differences between the absorbed and unabsorbed populations (e.g. Aird et al. 2015; Georgakakis et al. 2015). Clearly, the full picture of supermassive black holes (SMBHs) and active galactic nuclei (AGNs) evolving through cosmological time is complicated, and requires detailed investigation. Theoretical models and cosmological simulations have allowed us to try to quantify the role of different contributing black hole fuelling mechanisms (e.g. mergers, disc instabilities) and obscuration to the AGN luminosity function (e.g. Fanidakis et al. 2012; Hirschmann et al. 2012), but we do not yet fully understand the reasons for the different features of the evolution.

★ E-mail: andrew.j.griffin@durham.ac.uk

The evolution of AGNs through time also has significance for galaxy formation, since AGNs are thought to have a dramatic effect on their host galaxies. The relativistic jets from AGNs can have a strong effect on the surrounding hot gas by forming huge X-ray cavities (e.g. Forman et al. 2005; Blanton et al. 2011; Randall et al. 2011), or the AGNs can drive powerful high-velocity outflows (e.g. Pounds et al. 2003; Reeves, O'Brien & Ward 2003; Rupke & Veilleux 2011).

The precise physical mechanism for the production of AGN jets has not yet been determined, but the two most popular mechanisms are either that the accretion flow determines the jet power (Blandford & Payne 1982) or that the spin of the black hole determines the jet power (Blandford & Znajek 1977). Many simulations of black hole accretion discs have been conducted to study jet formation, where black hole spin often plays a key role (e.g. Kudoh, Matsumoto & Shibata 1998; Hawley & Balbus 2002; McKinney 2005; Hawley & Krolik 2006; Tchekhovskoy & McKinney 2012; Sądowski et al. 2013). The importance of black hole spin has motivated observational studies to constrain black hole spin values (Brenneman & Reynolds 2006; Chiang & Fabian 2011; Done et al. 2013) and cosmological simulations of black hole spin evolution (Berti & Volonteri 2008; Lagos, Padilla & Cora 2009; Fanidakis et al. 2011; Barausse 2012; Dotti et al. 2013; Fiacconi, Sijacki & Pringle 2018). The latter have been used to try to understand the role of black hole mergers and accretion across cosmological time on SMBH growth and spin evolution.

Constructing a model that accounts for all the processes involved in changing SMBH spin is not a simple task – especially given the vast range of scales involved. On sub-parsec scales, the spin of the black hole affects the radius of the last stable orbit for orbiting material, and hence the radiative efficiency of the black hole. The SMBH spin can be misaligned with that of the accretion disc, causing the accretion disc to become warped, which affects how gas is accreted (Bardeen & Petterson 1975). On larger scales, it is currently unclear whether the gas accretes in an ordered manner or in a series of randomly oriented events (King, Pringle & Hofmann 2008), which affects how much the black hole spins up. On galaxy-wide scales, simulations and theoretical models predict that the SMBH may accrete gas from cold, infalling gas made available by galaxy mergers (e.g. Di Matteo, Springel & Hernquist 2005) or by disc instabilities (e.g. Younger et al. 2008), or from gas accreted from the hot halo gas surrounding the galaxy (e.g. Bower et al. 2006). Cold streams may also supply gas to the central regions of galaxies (e.g. Khandai et al. 2012).

To be able to model all of these processes operating over all of these scales, some form of 'sub-grid' prescription is required to be able to model the effects going on below the numerical resolution of the calculation. Therefore, investigating SMBH spin evolution and mass growth and exploring their effects on galaxy wide scales is well suited to using a semi-analytic model of galaxy formation. Using a semi-analytic model of galaxy formation coupled with a large-volume, high-resolution dark matter simulation means that we can conduct detailed simulations within a computationally reasonable time-frame, which means that we can investigate the low-redshift Universe (by comparing to observations), and can make predictions for the high-redshift Universe, with greater accuracy than previous studies using simulations.

Semi-analytic models of galaxy formation have greatly contributed to our understanding of SMBHs and AGNs in galaxy formation. Bower et al. (2006) used GALFORM with an AGN feedback prescription in which heating of halo gas by relativistic jets balance radiative cooling in the most massive haloes to provide a match to the galaxy luminosity function at a range of redshifts, highlighting the potential importance of AGN feedback on galaxy formation. Malbon et al. (2007) extended the GALFORM model of Baugh et al. (2005) by including SMBH growth from mergers, cold gas accreted from starbursts and from the hot halo mode introduced in Bower et al. (2006) to reproduce the quasar optical luminosity function. In Fanidakis et al. (2011), GALFORM was updated to include an SMBH spin evolution model in which SMBH spin evolves during accretion of gas or by merging with other SMBHs. This model was then compared to observed AGN luminosity functions in the redshift range $0 < z < 6$ for optical and X-ray data in Fanidakis et al. (2012). Other semi-analytic galaxy formation models have also investigated SMBH growth and evolution (e.g. Lagos, Cora & Padilla 2008; Marulli et al. 2008; Bonoli et al. 2009; Hirschmann et al. 2012; Menci, Fiore & Lamastra 2013; Enoki et al. 2014; Neistein & Netzer 2014; Shirakata et al. 2019) and studies have also been conducted using hydrodynamical simulations (e.g. Hirschmann et al. 2014; Sijacki et al. 2015; Rosas-Guevara et al. 2016; Volonteri et al. 2016; Weinberger et al. 2018).

In this paper, we present predictions for the evolution of SMBH and AGN properties in the redshift range $0 < z < 6$, using an updated prescription for the evolution of SMBH spin within the GALFORM semi-analytic model of galaxy formation. We include a more detailed treatment of the obscuration and compare the model predictions to more recent observational data. In a subsequent paper we will present predictions for $z > 6$.

This paper is organized as follows. In Section 2, we outline the galaxy formation model and the spin evolution model. In Section 3, we outline the calculation of AGN luminosities. In Section 4, we present predictions for black hole masses and spins for the model, as well as the dependence of AGN luminosities on galaxy properties. In Section 5, we show the evolution of the AGN luminosity function at different wavelengths for $0 < z < 6$. In Section 6, we give concluding remarks.

2 THE GALAXY FORMATION AND SMBH EVOLUTION MODEL

2.1 The GALFORM model

To make our predictions, we use the Durham semi-analytic model of galaxy formation, GALFORM. Building on the principles outlined in White & Rees (1978), White & Frenk (1991), and Cole et al. (1994), and introduced in Cole et al. (2000), in GALFORM galaxies form from baryons condensing within dark matter haloes, with the assembly of the haloes described by the dark matter halo merger trees. While dark matter merger trees can be calculated using a Monte Carlo technique that is based on the Extended Press-Schechter theory (Lacey & Cole 1993; Cole et al. 2000; Parkinson, Cole & Helly 2008), they can also be extracted from dark matter *N*-body simulations (Kauffmann et al. 1999; Helly et al. 2003; Jiang et al. 2014), which is the method that we follow in this paper. The baryonic physics is then modelled using a set of coupled differential equations to track the exchange of baryons between different galaxy components. The physical processes modelled in GALFORM include (i) the merging of dark matter haloes, (ii) shock heating and radiative cooling of gas in haloes, (iii) star formation from cold gas, (iv) photoionization/supernova/AGN feedback, (v) the chemical evolution of gas and stars, (vi) galaxies merging in haloes due to dynamical friction, (vii) disc instabilities, (viii) the evolution of stellar populations, and (ix) the extinction and reprocessing of stellar radiation by dust. For a detailed description

of the physical processes involved, see Lacey et al. (2016) and references therein.

In this paper, we update the model for SMBHs and AGNs presented in Fanidakis et al. (2011), superceding the equations in that paper, which contained some typographical errors, and also putting special emphasis on improving the model for the obscuration of AGNs at X-ray and optical wavelengths. We incorporate the updated Fanidakis et al. (2011) SMBH model in the Lacey et al. (2016) GALFORM model as updated for the Planck–Millennium simulation by Baugh et al. (2018). The Lacey et al. (2016) model brings together several GALFORM developments into a single model, which fits well a wide range of observational data covering wavelengths from the far-UV to the sub-mm in the redshift range $0 < z < 6$. The Lacey et al. (2016) GALFORM model differs in a number of ways from that used in Fanidakis et al. (2011, 2012), including having different IMFs for quiescent and starburst star formation, as opposed to the single IMF used in Fanidakis et al. (2011).

The dark matter simulation used for these predictions is a new $(800\,\mathrm{Mpc})^3$ Millennium style simulation (Springel et al. 2005) with cosmological parameters consistent with the *Planck* satellite results (Planck Collaboration XVI 2014) – henceforth referred to as the P-Millennium (Baugh et al. 2018). The P-Millennium has an increased number of snapshots output – 270 instead of 64 for the Millennium simulation, the time interval between outputs has been chosen to ensure there are sufficient output snapshots for convergence of galaxy properties (cf. Benson et al. 2012). The halo mass resolution is $2.12 \times 10^9\,h^{-1}\,\mathrm{M}_\odot$, compared to the halo mass resolution of $1.87 \times 10^{10}\,h^{-1}\,\mathrm{M}_\odot$ for the dark matter simulation used in Lacey et al. (2016). This halo mass resolution is a result of P-Millennium having a dark matter particle mass of $1.06 \times 10^8\,h^{-1}\,\mathrm{M}_\odot$.

Because of the changed cosmological parameters and improved halo mass resolution in P-Millennium compared to the simulation used in Lacey et al. (2016), it was necessary to re-calibrate some of the galaxy formation parameters – which was done in Baugh et al. (2018). The new model also includes a more accurate calculation of the time-scale for galaxies to merge within a halo (Simha & Cole 2017). Only two GALFORM parameters were changed, both relating to supernova feedback and the return of ejected gas. The parameters were changed in Baugh et al. (2018) from the Lacey et al. (2016) model values are shown in Table 1. This P-Millennium based model has already been used in Cowley et al. (2018) to make predictions for galaxies for *JWST* in near- and mid-IR bands, and a model using P-Millennium and the model of Gonzalez-Perez et al. (2018) was used to study the effect of AGN feedback on halo occupation distribution models in McCullagh et al. (2017).

We emphasize that the aim in this paper is to study SMBH and AGN evolution in the framework of an existing galaxy formation model calibrated on a wide range observational data on galaxies. Therefore, we do not consider any modifications to the underlying galaxy formation model, only to the modelling of SMBH and AGN within it.

2.2 SMBH growth

SMBHs in GALFORM grow in three different ways.

2.2.1 *Starburst mode gas accretion*

First, SMBHs can accrete gas during starbursts, which are triggered by either galaxy mergers or disc instabilities. In both of these cases, all of the remaining cold gas in a galaxy is consumed in a starburst and a fixed fraction of the mass of stars formed from the starburst feeds the SMBH, such that the accreted mass $M_{\rm acc}$ is $f_{\rm BH}M_{\star,\rm burst}$ where $M_{\star,\rm burst}$ is the mass of stars formed in the starburst and $f_{\rm BH}$ is a free parameter (cf. Lacey et al. 2016). Note that the mass of the stars formed is less than the initial mass of the gas in the starburst due to the ejection of gas by supernova feedback.

A galaxy merger can cause gas to be transferred to the centre of the galaxy and trigger a burst of star formation (e.g. Mihos & Hernquist 1996). Some of this gas is then available to feed the central SMBH (Kauffmann & Haehnelt 2000; Malbon et al. 2007). In the model, if the mass ratio of the two galaxies is less than 0.3, the merger is defined as a minor merger, and a starburst is triggered for a minor merger with mass ratio above 0.05. If the mass ratio of the galaxies is greater than 0.3, the merger is defined as a major merger, and a spheroid is formed. This is described in section 3.6.1 of Lacey et al. (2016).

Disc instabilities cause a bar to be formed, which disrupts the galaxy disc (Efstathiou, Lake & Negroponte 1982) and transfers gas to the centre of the galaxy to be fed into the SMBH. Disc instabilities driving gas into the centres of galaxies is an effect seen in various hydrodynamical simulations (e.g. Hohl 1971; Bournaud, Combes & Semelin 2005; Younger et al. 2008), and used as a channel of black hole/bulge growth in many semi-analytic models of galaxy formation (e.g. De Lucia et al. 2011; Hirschmann et al. 2012; Menci et al. 2014; Croton et al. 2016; Lagos et al. 2018), although the implementation of these disc instabilities varies between models. Most models use the disc instability criterion of Efstathiou et al. (1982) in which the disc becomes unstable if it is sufficiently self-gravitating; however, different models use this condition differently. For example, in the model of Hirschmann et al. (2012), if a disc is unstable, then enough gas and stars are transferred from the disc to the bulge to completely stabilize the disc, while in GALFORM, we assume that if a disc is unstable, then it is completely destroyed and forms a bulge. Numerical simulations of isolated discs show that disc instabilities can transfer large fractions of gas and stars into the bulge in some situations (e.g. Bournaud, Elmegreen & Elmegreen 2007; Elmegreen, Bournaud & Elmegreen 2008; Saha & Cortesi 2018). Disc instabilities in GALFORM are described in section 3.6.2 of Lacey et al. (2016).

2.2.2 *Hot halo mode gas accretion*

In GALFORM, we assume that SMBHs can also accrete gas from the hot gas atmospheres of massive haloes: when large haloes collapse, gas is shock heated to form a quasistatic hot halo atmosphere. For sufficiently massive haloes, the cooling time of this gas is longer than its free-fall time, and the SMBH is fed with a slow inflow from the halo's hot atmosphere – 'hot halo mode accretion' (Bower et al. 2006). The black hole is assumed to grow by this fuelling mechanism only when AGN feedback is operational. In this regime, energy input by a relativistic jet is assumed to balance radiative cooling in the halo, with the mass accretion rate on to the black hole $\dot{M}$ being determined by this energy balance condition. The mass $M_{\rm acc}$ accreted on to the SMBH in a simulation time-step $\Delta t_{\rm step}$ is then $\dot{M}\,\Delta t_{\rm step}$. The hot halo accretion mode is fully described in section 3.5.3 of Lacey et al. (2016).

2.2.3 *SMBH mergers*

SMBHs can be built up by SMBH–SMBH mergers. When galaxies merge, dynamical friction from gas, stars and dark matter causes the SMBH of the smaller galaxy to sink towards the other SMBH.

Table 1. The cosmological and galaxy formation parameters for this study that have been changed from the Lacey et al. (2016) GALFORM model. These parameters in this study are as in Baugh et al. (2018). Ω_{m0}, Ω_{v0}, and Ω_{b0} are the present-day density parameters in matter, vacuum energy, and baryons; h is the present-day Hubble parameter in units of 100 km s^{-1} Mpc^{-1} and σ_8 is the normalization of the initial power spectrum of density fluctuations. The parameters γ_{SN} and α_{ret} are related to supernova feedback and the return time-scale for ejected gas, as described in section 3.5.2 in Lacey et al. (2016).

Parameter	Description	Lacey et al. (2016)	This study
Ω_{m0}	Matter density	0.272	0.307
Ω_{v0}	Vacuum energy density	0.728	0.693
h	Reduced Hubble parameter	0.704	0.678
Ω_{b0}	Baryon density	0.0455	0.0483
σ_8	Power spectrum normalization	0.818	0.829
α_{ret}	Gas reincorporation time-scale	0.64	1.0
γ_{SN}	Slope of SN feedback mass loading	3.2	3.4
	Galaxy merger time-scale	Jiang et al. (2008)	Simha & Cole (2017)

Then, as the separation decreases, gravitational radiation provides a mechanism by which the SMBHs can lose angular momentum and spiral in to merge and form a larger SMBH. In the model, we assume the time-scale on which the SMBHs merge is short, so that the SMBHs merge when the galaxies merge.

2.3 SMBH seeds

The starting point for the treatment of SMBHs in the model is SMBH seeds that eventually grow by accretion of gas and by merging with other SMBHs to form the objects in the Universe today. The processes for SMBH seed formation are uncertain (see e.g. Volonteri 2010, and references therein) and so we simply add a seed SMBH of mass M_{seed} into each halo, where M_{seed} is a parameter that we can vary. Unless otherwise stated, this parameter has the value $M_{seed} = 10\,h^{-1}\,M_\odot$ – representative of the SMBH seeds formed by stellar collapse. The effect of varying this seed mass is discussed in Appendix A.

2.4 SMBH mass growth and spin-up by gas accretion

Our model includes the evolution of SMBH spin. In this model, SMBHs can change spin in two ways: (i) by accretion of gas or (ii) by merging with another SMBH. The SMBH spin is characterized by the dimensionless spin parameter, $a = cJ_{BH}/GM_{BH}^2$, within the range $-1 \le a \le 1$, where J_{BH} is the angular momentum of the SMBH, and M_{BH} is the mass of the SMBH. $a = 0$ represents a black hole that is not spinning and $a = 1$ or $a = -1$ represents a maximally spinning black hole. The sign of a is defined by the direction of the angular momentum of the black hole relative to that of the innermost part of the accretion disc, so for $a > 0$ the black hole is spinning in the same direction as the inner accretion disc and for $a < 0$ the black hole is spinning in the opposite direction to the inner accretion disc. To calculate the SMBH spin, a^f after an accretion episode, we use the expression in Bardeen (1970):[1]

$$a^f = \frac{1}{3}\sqrt{\hat{r}_{lso,i}}\frac{M_{BH,i}}{M_{BH,f}}\left(4 - \left[3\hat{r}_{lso,i}\left(\frac{M_{BH,i}}{M_{BH,f}}\right)^2 - 2\right]^{1/2}\right), \quad (1)$$

where $\hat{r}_{lso}$ is the radius of the last stable circular orbit in units of the gravitational radius, $R_G = GM_{BH}/c^2$, and the subscripts i and f indicate values at the start and end of an accretion event. The black hole mass before and after an accretion event are related by

$$M_{BH,f} = M_{BH,i} + (1 - \epsilon_{TD})\Delta M, \quad (2)$$

where ΔM is the mass accreted from the disc in this accretion episode and ϵ_{TD}, the radiative accretion efficiency for a thin accretion disc, is given by

$$\epsilon_{TD} = 1 - \left(1 - \frac{2}{3\hat{r}_{lso}}\right)^{1/2}. \quad (3)$$

$\hat{r}_{lso}$ is calculated from the spin a, as in Bardeen, Press & Teukolsky (1972):

$$\hat{r}_{lso} = 3 + Z_2 \mp \sqrt{(3 - Z_1)(3 + Z_1 + 2Z_2)}, \quad (4)$$

with the minus sign for $a > 0$ and the positive sign for $a < 0$. The functions Z_1 and Z_2 are given by

$$Z_1 = 1 + (1 - |a|^2)^{1/3}[(1 + |a|)^{1/3} + (1 - |a|)^{1/3}], \quad (5)$$

$$Z_2 = \sqrt{3|a|^2 + Z_1^2}. \quad (6)$$

We consider the accretion disc in three separate parts as shown in Fig. 1 – an outer disc at radii greater than an inner radius, R_{in}, an inner disc for radii less than R_{in}, and a warped disc for radii less than the warp radius, R_{warp}. The SMBH has an angular momentum $\boldsymbol{J}_{BH}$, and the angular momentum of the disc within R_{in} is $\boldsymbol{J}_{in}$. If $\boldsymbol{J}_{BH}$ is not in the same direction as $\boldsymbol{J}_{in}$ a spinning black hole induces a Lense–Thirring precession in the misaligned disc elements. Because the precession rate falls off as R^{-3}, at smaller radii the black hole angular momentum and the accretion disc angular momentum vectors will become exactly aligned or anti-aligned, whereas at sufficiently large radii there will still be a misalignment (Bardeen & Petterson 1975). The transition between these two regions occurs at the so-called warp radius, R_{warp}. The angular momentum of the disc within the warp radius is $\boldsymbol{J}_{warp}$. At the start of an accretion event, the angular momentum $\boldsymbol{J}_{warp}$ within R_{warp} is assumed to be aligned with $\boldsymbol{J}_{in}$. As a result of the torques, $\boldsymbol{J}_{BH}$ then aligns with $\boldsymbol{J}_{tot} = \boldsymbol{J}_{BH} + \boldsymbol{J}_{warp}$ (which remains constant during this alignment process) and $\boldsymbol{J}_{warp}$ either anti-aligns or aligns with $\boldsymbol{J}_{BH}$ (King et al. 2005). The gas within R_{warp} is then assumed to be accreted on to the SMBH from the aligned/anti-aligned disc. As more gas is accreted, $\boldsymbol{J}_{BH}$ eventually aligns with the rest of the inner disc, as the gas in the inner disc is consumed.

We consider two alternative scenarios for how the angular momentum directions of the inner and outer discs are related. In

[1] Note that equation (1) is corrected from Fanidakis et al. (2011, equation 6).

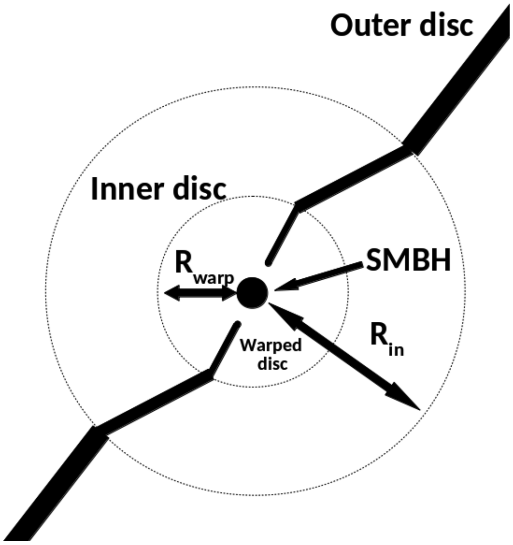

Figure 1. A diagram showing the various scales involved in the gas accretion – the warp radius $R_{\rm warp}$ and the inner radius $R_{\rm in}$. We refer to the region within $R_{\rm in}$ as the inner disc and the region outside of $R_{\rm in}$ as the outer disc.

the 'prolonged mode' accretion scenario, the angular momentum of the inner disc is in the same direction as the angular momentum of the outer disc, $\boldsymbol{J}_{\rm out}$, but in the 'chaotic mode' accretion scenario introduced in King et al. (2008); the orientation of the angular momentum of the inner disc is randomly oriented with respect to the angular momentum of the outer disc. King et al. (2008) propose that $R_{\rm in}$ is the self-gravity radius of the disc, and we assume this in our model.

The motivation for chaotic mode accretion is twofold. First, the Soltan (1982) argument, a comparison of the integral of the quasar luminosity function over luminosity and redshift to the integral over the black hole mass function in the local Universe, implies an average radiative efficiency of SMBH growth of $\epsilon \approx 0.1$ (which corresponds to a spin value of $a \approx 0.67$), suggesting that SMBHs in the Universe are typically not maximally spinning, as we would expect from SMBHs that have been spun up by the accretion of gas that is aligned in the same direction, as in the prolonged accretion scenario. Secondly, AGN jets seem to be misaligned with their host galaxies (e.g. Kinney et al. 2000; Sajina et al. 2007), suggesting a misaligned accretion of material on to the SMBH.

Accretion continues in this manner until the gas in the outer disc has been consumed. For this analysis, we adopt chaotic mode accretion as our standard choice.

2.5 Warped accretion discs

To obtain the warp radius, $R_{\rm warp}$, of an accretion disc, we need expressions for the structure of the accretion disc. There are two different types of accretion discs: (i) physically thin, optically thick, radiatively efficient 'thin discs' (Shakura & Sunyaev 1973) and (ii) physically thick, optically thin, radiatively inefficient advection-dominated accretion flows (ADAFs; see Yuan & Narayan 2014 for a review). Shakura & Sunyaev (1973) introduced the 'α-prescription' to solve the accretion disc equations for a thin disc, where the viscosity, ν, is given by $\nu = \alpha_{\rm TD} c_{\rm s} H$, where $\alpha_{\rm TD}$ is the dimensionless Shakura & Sunyaev (1973) parameter, $c_{\rm s}$ is the sound speed and H is the disc semi-thickness. In this analysis, we use the solutions of Collin-Souffrin & Dumont (1990), in which the accretion disc equations are solved for AGN discs, assuming this α-prescription. We use their solution for the regime where the opacity is dominated by electron scattering and where gas pressure dominates over radiation pressure.

The disc surface density, Σ, is then given by

$$\Sigma = 6.84 \times 10^5 \,{\rm g\,cm^{-2}}\, \alpha_{\rm TD}^{-4/5} \dot{m}^{3/5} \left(\frac{M_{\rm BH}}{10^8\,{\rm M_\odot}}\right)^{1/8} \left(\frac{R}{R_{\rm S}}\right)^{-3/5}, \quad (7)$$

where $\dot{m} = \dot{M}/\dot{M}_{\rm Edd}$ is the dimensionless mass accretion rate, R is the radius from the centre of the disc, and $R_{\rm S} = 2GM_{\rm BH}/c^2$ is the Schwarzschild radius. The value we use for $\alpha_{\rm TD}$ is given in Table 2. The disc semi-thickness H is given by[2]

$$\frac{H}{R} = 1.25 \times 10^{-3} \alpha_{\rm TD}^{-1/10} \dot{m}^{1/5} \left(\frac{M_{\rm BH}}{10^8\,{\rm M_\odot}}\right)^{-1/10} \left(\frac{R}{R_{\rm S}}\right)^{1/20}. \quad (8)$$

We calculate the Eddington luminosity using

$$L_{\rm Edd} = \frac{4\pi G M_{\rm BH} c}{\kappa} = 1.26 \times 10^{46} \left(\frac{M_{\rm BH}}{10^8\,{\rm M_\odot}}\right) {\rm erg\,s^{-1}}, \quad (9)$$

where κ is the opacity, for which we have used the electron scattering opacity for pure hydrogen gas. We calculate the Eddington mass accretion rate $\dot{M}_{\rm Edd}$ from $L_{\rm Edd}$ using a nominal accretion efficiency $\epsilon = 0.1$ (as used in Yuan & Narayan 2014) chosen so that the Eddington normalized mass accretion rate $\dot{m}$ does not depend on the black hole spin:

$$\dot{M}_{\rm Edd} = \frac{L_{\rm Edd}}{0.1c^2}. \quad (10)$$

Note that for the calculation of the luminosities, we do use the spin-dependent radiative efficiency. We then follow the method of Natarajan & Pringle (1998) and Volonteri, Sikora & Lasota (2007) and take the warp radius as the radius at which the time-scale for radial diffusion of the warp due to viscosity is equal to the local Lense–Thirring precession time-scale. This then gives an expression for the warp radius:[3]

$$\frac{R_{\rm warp}}{R_S} = 3410\, a^{5/8} \alpha_{\rm TD}^{-1/2} \dot{m}^{-1/4} \left(\frac{M_{\rm BH}}{10^8\,{\rm M_\odot}}\right)^{1/8} \left(\frac{\nu_2}{\nu_1}\right)^{-5/8}, \quad (11)$$

where $\nu_{1,2}$ are the horizontal and vertical viscosities, respectively. For this analysis, we assume that $\nu_1 = \nu_2$ (e.g. King et al. 2008). The warp mass can then be calculated using

$$M_{\rm warp} = \int_0^{R_{\rm warp}} 2\pi \Sigma(R) R^2 {\rm d}R, \quad (12)$$

to give an expression[4]

$$M_{\rm warp} = 1.35\ {\rm M_\odot} \alpha_{\rm TD}^{-4/5} \dot{m}^{3/5} \left(\frac{M_{\rm BH}}{10^8\,{\rm M_\odot}}\right)^{11/5} \left(\frac{R_{\rm warp}}{R_{\rm S}}\right)^{7/5}. \quad (13)$$

2.6 Self-gravitating discs

In the chaotic mode accretion scenario of King et al. (2008), the inner radius, $R_{\rm in}$, is assumed to be equal to the disc self-gravity radius, $R_{\rm sg}$. The self-gravity radius of the accretion disc is the radius at which the vertical gravity due to the disc equals the vertical gravity of the central SMBH at the disc mid-plane. For thin discs (where

[2]Note that equation (8) is different to Fanidakis et al. (2011, equation 25).
[3]Note that equation (11) is different to Fanidakis et al. (2011, equation 15).
[4]Note that equation (13) is different to Fanidakis et al. (2011, equation 18).

Table 2. The values for the SMBH/AGN free parameters in the model. The upper part of the table shows parameters where the values adopted are from other studies, whereas the lower part of the table gives parameters that have been calibrated on the luminosity functions in Section 4.4.

Parameter	Fanidakis et al. (2012)	Adopted here	Significance
$\alpha_{\rm ADAF}$	0.087	0.1	Shakura & Sunyaev (1973) viscosity parameter for ADAFs
$\alpha_{\rm TD}$	0.087	0.1	Shakura & Sunyaev (1973) viscosity parameter for TDs
$\delta_{\rm ADAF}$	2000^{-1}	0.2	Fraction of viscous energy transferred to electrons in ADAF
$\dot{m}_{\rm crit,ADAF}$	0.01	0.01	Boundary between thin disc and ADAF accretion
$\eta_{\rm Edd}$	4	4	Super-Eddington suppression factor
$f_{\rm q}$	10	10	Ratio of lifetime of AGN episode to bulge dynamical time-scale

$\dot{m} > \dot{m}_{\rm crit,ADAF}$), the self-gravity condition is (Pringle 1981)

$$M_{\rm sg} = M_{\rm BH}\frac{H}{R}, \tag{14}$$

where $M_{\rm sg}$ is the disc mass within the radius $R_{\rm sg}$. For ADAFs (where $\dot{m} < \dot{m}_{\rm crit,ADAF}$), $H \sim R$, so the self-gravity condition is

$$M_{\rm sg} = M_{\rm BH}. \tag{15}$$

Using the accretion disc solutions of Collin-Souffrin & Dumont (1990), we derive an expression for the self-gravity radius for thin discs:[5]

$$\frac{R_{\rm sg}}{R_{\rm S}} = 4790\,\alpha_{\rm TD}^{14/27}\dot{m}^{-8/27}\left(\frac{M_{\rm BH}}{10^8\,{\rm M}_\odot}\right)^{-26/27}, \tag{16}$$

and using an integral similar to equation (12), the self-gravity mass for the thin disc is given by[6]

$$M_{\rm sg} = 1.35\ {\rm M}_\odot\alpha_{\rm TD}^{-4/5}\dot{m}^{3/5}\left(\frac{M_{\rm BH}}{10^8\,{\rm M}_\odot}\right)^{11/5}\left(\frac{R_{\rm sg}}{R_{\rm S}}\right)^{7/5}. \tag{17}$$

2.7 Numerical procedure for modelling SMBH accretion

We have calculated results for both the prolonged and chaotic scenario, and for gas accreted in increments of the self-gravity mass or warp mass. We present predictions mostly for our standard case in which mass is accreted in increments of the self-gravity mass and assuming the chaotic mode of accretion. We find that the predicted spin distribution of the SMBHs is the same if we use increments of the self-gravity mass or the warp mass (cf. Fig. 8) and so we use increments of the self-gravity mass as it is computationally faster. This is because when gas is accreted on to the SMBH in increments of the warp mass, for small SMBHs the warp mass is very small, and so in each accretion event the SMBH grows by a very small amount in each accretion event. First, we present the numerical procedure when mass is accreted in increments of the warp mass (cf. Volonteri et al. 2007; Fanidakis et al. 2011), and then the case where mass is accreted in increments of the self-gravity mass (cf. King et al. 2008).

2.7.1 *Accretion in increments of the warp mass*

For the first warp mass of gas, the angular momentum of the SMBH, $\boldsymbol{J}_{\rm BH}$, and the angular momentum of the inner disc, $\boldsymbol{J}_{\rm in}$, are assigned a random angle, θ_i, in the range $[0, \pi]$ radians. In the chaotic mode, each time the inner disc is consumed, θ_i, is assigned a new random angle. The gas with $R < R_{\rm warp}$ initially has angular momentum $\boldsymbol{J}_{\rm warp}$ aligned with $\boldsymbol{J}_{\rm in}$, so θ_i is also the initial angle between $\boldsymbol{J}_{\rm BH}$ and $\boldsymbol{J}_{\rm warp}$.

[5]Note that equation (16) is different to Fanidakis et al. (2011), equation 24).
[6]Note that equation (17) is different to Fanidakis et al. (2011), equation 26).

$\boldsymbol{J}_{\rm BH}$ and $\boldsymbol{J}_{\rm warp}$ are then evolved according to the Lense–Thirring effect described in Section 2.4, with $\boldsymbol{J}_{\rm BH}$ and $\boldsymbol{J}_{\rm warp}$, respectively, aligning and aligning/anti-aligning with $\boldsymbol{J}_{\rm tot}$. The magnitude of $\boldsymbol{J}_{\rm BH}$ remains constant during this process, but the magnitude of $\boldsymbol{J}_{\rm warp}$ changes. This is treated as happening before the mass consumption on to the SMBH starts.

We calculate the angular momentum of the material within the warped disc as $J_{\rm warp} = M_{\rm warp}\sqrt{GM_{\rm BH}R_{\rm warp}}$ and the angular momentum of the black hole, $J_{\rm BH} = 2^{-1/2}M_{\rm BH}a\sqrt{GM_{\rm BH}R_{\rm S}}$. Then the ratio of these two quantities is

$$\frac{J_{\rm warp}}{2J_{\rm BH}} = \frac{M_{\rm warp}}{\sqrt{2}aM_{\rm BH}}\left(\frac{R_{\rm warp}}{R_{\rm S}}\right)^{1/2}. \tag{18}$$

Whether $\boldsymbol{J}_{\rm warp}$ and $\boldsymbol{J}_{\rm BH}$ align or anti-align with each other depends on this ratio and on the angle θ_i. If $\cos\theta_i > -J_{\rm warp}/2J_{\rm BH}$, $\boldsymbol{J}_{\rm warp}$ and $\boldsymbol{J}_{\rm BH}$ become aligned (prograde accretion), whereas if $\cos\theta_i < -J_{\rm warp}/2J_{\rm BH}$, $\boldsymbol{J}_{\rm warp}$ and $\boldsymbol{J}_{\rm BH}$ become anti-aligned (retrograde accretion). The angle between $\boldsymbol{J}_{\rm BH}$ and $\boldsymbol{J}_{\rm in}$ after the accretion event, θ_f, is determined by conservation of $\boldsymbol{J}_{\rm tot}$ and $|\boldsymbol{J}_{\rm BH}|$ and is given by

$$\cos\theta_f = \frac{J_{\rm warp} + J_{\rm BH}\cos\theta_i}{\sqrt{J_{\rm BH}^2 + J_{\rm warp}^2 + 2J_{\rm warp}J_{\rm BH}\cos\theta_i}}. \tag{19}$$

When a new warp mass $M_{\rm warp}$ is then consumed, the gas is given a new $\boldsymbol{J}_{\rm warp}$ pointing in the same direction as the inner disc and the same process happens again. This repeated process has the effect that $\boldsymbol{J}_{\rm BH}$ gradually aligns with the angular momentum of the inner accretion disc, $\boldsymbol{J}_{\rm in}$, as more gas is accreted. Eventually, the gas in the inner disc is completely consumed.

In the prolonged mode, this process continues until all of the gas in the outer disc has also been consumed, whereas in the chaotic mode, once a self-gravity mass of gas has been consumed, the angle between $\boldsymbol{J}_{\rm in}$ and $\boldsymbol{J}_{\rm out}$ is randomized again.

2.7.2 *Accretion in increments of the self-gravity mass*

In the scenario where gas is being accreted in increments of the self-gravity mass of gas, the above procedure is followed, but only once for each inner disc of gas consumed. For this case, the ratio of angular momenta is given by

$$\frac{J_{\rm in}}{2J_{\rm BH}} = \frac{M_{\rm sg}}{\sqrt{2}aM_{\rm BH}}\left(\frac{\min(R_{\rm warp}, R_{\rm sg})}{R_{\rm S}}\right)^{1/2}. \tag{20}$$

In the future, we plan a more thorough analysis of the effect on the spin evolution of accreting in increments of self-gravity mass compared to increments of warp mass. The AGN luminosities are not affected by this choice as they depend on the accreted mass and the SMBH spin as we describe in Section 3.1.

2.8 Spin-up by SMBH mergers

The other way in which an SMBH can change its spin is by merging with another SMBH. The spin of the resulting SMBH depends on the spins of the two SMBHs that merge and on the angular momentum of their binary orbit. To determine the final spin, $\boldsymbol{a}_f$, we use the expressions obtained from numerical simulations of BH–BH mergers in Rezzolla et al. (2008):

$$|\boldsymbol{a}_f| = \frac{1}{(1+q)^2}\left(|\boldsymbol{a}_1^2| + |\boldsymbol{a}_2^2|q^4 + 2|\boldsymbol{a}_1||\boldsymbol{a}_2|q^2\cos\phi + 2\left(|\boldsymbol{a}_1|\cos\theta + |\boldsymbol{a}_2|q^2\cos\xi\right)|\boldsymbol{l}|q + |\boldsymbol{l}|^2q^2\right)^{1/2}, \quad (21)$$

where $\boldsymbol{a}_{1,2}$ are the spins of the SMBHs, q is the mass ratio M_1/M_2, with M_1 and M_2 chosen such that $q \leq 1$, μ is the symmetric mass ratio $q/(q+1)^2$, and $\boldsymbol{l}$ is the contribution of the orbital angular momentum to the spin angular momentum of the final black hole. It is assumed that the direction of $\boldsymbol{l}$ is that of the initial orbital angular momentum, while its magnitude is given by

$$|\boldsymbol{l}| = \frac{s_4}{(1+q^2)^2}\left(|\boldsymbol{a}_1|^2 + |\boldsymbol{a}|_1^2q^4 + 2|\boldsymbol{a}_1||\boldsymbol{a}_2|q^2\cos\phi\right) + \left(\frac{s_5\mu + t_0 + 2}{1+q^2}\right)\left(|\boldsymbol{a}_1|\cos\theta + |\boldsymbol{a}_2|q^2\cos\xi\right) + 2\sqrt{3} + t_2\mu + t_3\mu^2. \quad (22)$$

where $s_4 = -0.129$, $s_5 = -0.384$, $t_0 = -2.686$, $t_2 = -3.454$, $t_3 = 2.353$ are values obtained in Rezzolla et al. (2008). The angles ϕ, θ, and ξ are the angles between the spins of the two black holes and their orbital angular momentum, and are given by

$$\cos\phi = \hat{\boldsymbol{a}}_1 \cdot \hat{\boldsymbol{a}}_2, \quad (23)$$

$$\cos\theta = \hat{\boldsymbol{a}}_1 \cdot \hat{\boldsymbol{l}}, \quad (24)$$

$$\cos\xi = \hat{\boldsymbol{a}}_2 \cdot \hat{\boldsymbol{l}}. \quad (25)$$

When we consider two SMBHs merging, we calculate the angles between the three different vectors by randomly selecting directions for $\boldsymbol{a_1}$, $\boldsymbol{a_2}$, and $\boldsymbol{l}$ uniformly over the surface of a sphere. This prescription makes the assumption that the radiation of gravitational waves does not affect the direction of the orbital angular momentum as the binary orbit shrinks, and we also assume that the mass lost to gravitational radiation is negligible.

3 CALCULATING AGN LUMINOSITIES

3.1 AGN bolometric luminosities

From the mass of gas that is accreted on to the SMBH, we can calculate a radiative bolometric luminosity as follows. In the starburst mode, we assume that during an accretion episode the accretion rate is constant over a time $f_q t_{\rm bulge}$, where $t_{\rm bulge}$ is the dynamical time-scale of the bulge and f_q is a free parameter, given in Table 2. Therefore, the mass accretion rate is given by

$$\dot{M} = \frac{M_{\rm acc}}{f_q t_{\rm bulge}}, \quad (26)$$

where $M_{\rm acc}$ is as defined in Section 2.2. In the hot halo mode, which is only active when AGN feedback is active, the mass accretion rate is determined by the condition that the energy released is just enough to balance radiative cooling:

$$\dot{M} = \frac{L_{\rm cool}}{\epsilon_{\rm heat}c^2}, \quad (27)$$

where $L_{\rm cool}$ is the radiative cooling luminosity of the hot halo gas, and $\epsilon_{\rm heat}$ is the efficiency of halo heating, which is treated as a free parameter (cf. Lacey et al. 2016).

We then calculate the bolometric luminosity for a thin accretion disc using

$$L_{\rm bol,TD} = \epsilon_{\rm TD}\dot{M}c^2, \quad (28)$$

where the radiative efficiency $\epsilon_{\rm TD}$ for the thin disc case depends on the black hole spin, as given by equation (3). However, the radiative efficiency is not the same for all regimes of the accretion flow. As well as the thin disc and the ADAF case, there are also AGNs accreting above the Eddington accretion rate. Such objects are generally understood to be advection dominated and to have optically thick flows (Abramowicz et al. 1988).

For the ADAF regime, we use the expressions for bolometric luminosity from Mahadevan (1997). There are two cases within this regime. For lower accretion rate ADAFs ($\dot{m} < \dot{m}_{\rm crit,visc}$), heating of the electrons is dominated by viscous heating, whereas for higher accretion rate ADAFs ($\dot{m}_{\rm crit,visc} < \dot{m} < \dot{m}_{\rm crit,ADAF}$), the ion–electron heating dominates the heating of the electrons. In the super-Eddington regime, the radiative efficiency is lower than the corresponding thin disc radiative efficiency, and so a super-Eddington luminosity suppression is introduced (Shakura & Sunyaev 1973). This expression includes a free parameter, $\eta_{\rm Edd}$, the value for which is given in Table 2.

Hence, the bolometric luminosities in the model are given by the following expressions.[7] For the low accretion rate ADAF regime, where $\dot{m} < \dot{m}_{\rm crit,visc}$:

$$L_{\rm bol} = 0.0002\epsilon_{\rm TD}\dot{M}c^2\left(\frac{\delta_{\rm ADAF}}{0.0005}\right)\left(\frac{1-\beta}{0.5}\right)\left(\frac{6}{\hat{r}_{\rm lso}}\right). \quad (29)$$

For the higher accretion rate ADAF regime, where $\dot{m}_{\rm crit,visc} < \dot{m} < \dot{m}_{\rm crit,ADAF}$, we have

$$L_{\rm bol} = 0.2\epsilon_{\rm TD}\dot{M}c^2\left(\frac{\dot{m}}{\alpha_{\rm ADAF}^2}\right)\left(\frac{\beta}{0.5}\right)\left(\frac{6}{\hat{r}_{\rm lso}}\right). \quad (30)$$

For the thin disc regime, where $\dot{m}_{\rm crit,ADAF} < \dot{m} < \eta_{\rm Edd}$, $L_{\rm bol} = L_{\rm bol,TD}$. Finally, for the super-Eddington regime, where $\dot{m} > \eta_{\rm Edd}$, we have

$$L_{\rm bol} = \eta_{\rm Edd}(1 + \ln(\dot{m}/\eta_{\rm Edd}))L_{\rm Edd}. \quad (31)$$

The value of $\eta_{\rm Edd}$ adopted gives a similar luminosity at a given mass accretion rate in the super-Eddington regime to the model of Watarai et al. (2000) who model super-Eddington sources as advection dominated slim discs.

In the above, $\alpha_{\rm ADAF}$ is the viscosity parameter in the ADAF regime (the value is given in Table 2). $\delta_{\rm ADAF}$ is the fraction of viscous energy transferred to the electrons (the value is given in Table 2). The current consensus for the value of $\delta_{\rm ADAF}$ is a value between 0.1 and 0.5 (cf. Yuan & Narayan 2014). Therefore, for this study, we adopt a value $\delta_{\rm ADAF} = 0.2$, more in line with observational (Yuan, Quataert & Narayan 2003; Liu & Wu 2013) and theoretical (Sharma et al. 2007) constraints, as opposed to the value of $\delta_{\rm ADAF} = 2000^{-1}$

[7]Note that the coefficients of the ADAF luminosities are derived in Mahadevan (1997) and not free parameters.

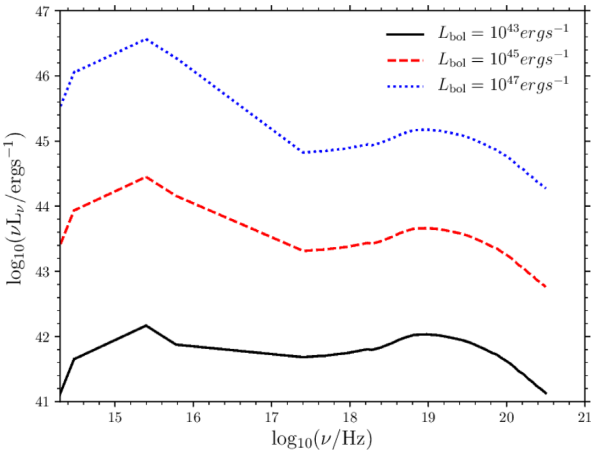

Figure 2. The Marconi et al. (2004) SED used for calculating luminosities in different wavebands in this work. Shown is the SED for $L_{\rm bol} = 10^{43}$ erg s^{-1} (black solid line), for $L_{\rm bol} = 10^{45}$ erg s^{-1} (red dashed line), and for $L_{\rm bol} = 10^{47}$ erg s^{-1} (blue dotted line).

adopted in Fanidakis et al. (2012). Changing the value of $\delta_{\rm ADAF}$ makes no discernible difference to the luminosity functions shown in this paper. β is the ratio of gas pressure to total pressure (total pressure being the sum of gas pressure and magnetic pressure). Following Fanidakis et al. (2012), we use the relation $\beta = 1 - \alpha_{\rm ADAF}/0.55$, which is based on MHD simulations in Hawley, Gammie & Balbus (1995).

The boundary between the two ADAF regimes is

$$\dot{m}_{\rm crit,visc} = 0.001 \left(\frac{\delta_{\rm ADAF}}{0.0005}\right) \left(\frac{1-\beta}{\beta}\right) \alpha^2_{\rm ADAF}, \quad (32)$$

which is a value chosen so that $L_{\rm bol}$ is continuous in the ADAF regime. The boundary between the ADAF and thin disc regimes is assumed to be $\dot{m}_{\rm crit,ADAF} = 0.01$ (Yuan & Narayan 2014). $f_{\rm q}$ and $\eta_{\rm Edd}$ are free parameters that we calibrate on observed AGN luminosity functions, as described in Section 5.1.

3.2 Converting from bolometric to optical and X-ray AGN luminosities

To convert from AGN bolometric luminosity to luminosities in other wavebands we use bolometric corrections derived from the empirical AGN SED template in Marconi et al. (2004). We show this SED for three different luminosities in Fig. 2. The rest-frame bolometric corrections calculated from this SED are[8]

$$\log_{10}(L_{\rm HX}/L_{\rm bol}) = -1.54 - 0.24\mathcal{L} - 0.012\mathcal{L}^2 + 0.0015\mathcal{L}^3, \quad (33)$$

$$\log_{10}(L_{\rm SX}/L_{\rm bol}) = -1.65 - 0.22\mathcal{L} - 0.012\mathcal{L}^2 + 0.0015\mathcal{L}^3, \quad (34)$$

$$\log_{10}(\nu_B L_{\nu B}/L_{\rm bol}) = -0.80 + 0.067\mathcal{L} - 0.017\mathcal{L}^2 + 0.0023\mathcal{L}^3, \quad (35)$$

where $\mathcal{L} = \log_{10}(L_{\rm bol}/10^{12} L_\odot)$, $L_{\rm HX}$ is the hard X-ray (2–10 keV) luminosity, $L_{\rm SX}$ is the soft X-ray (0.5–2 keV) luminosity, $\nu_B = c/4400$ Å is the frequency of the centre of the B band, and $L_{\nu B}$ is the luminosity per unit frequency in the B band.

To calculate B-band magnitudes we use the expression:[9]

$$M_{B,AB} = -11.33 - 2.5\log_{10}\left(\frac{\nu_B L_{\nu B}}{10^{40}\,{\rm ergs}^{-1}}\right), \quad (36)$$

for magnitudes in the AB system, from the definition of AB magnitudes (Oke & Gunn 1983). Using the Marconi et al. (2004) SED template, we convert from rest-frame B-band magnitudes to rest frame 1500 Å band magnitudes using a relation similar to equation (B8) to give

$$M_{1500,AB} = M_{B,AB} + 0.514. \quad (37)$$

The Marconi et al. (2004) SED is based on observations of quasars, with the UV part of the SED based on observations at $L_{\rm UV} \sim 10^{42.5-47}$ ergs^{-1} and the X-ray part of the SED based on observations at $L_{HX} \sim 10^{41-44}$ erg s^{-1}. Therefore, this SED is likely to be most appropriate for AGNs in the thin disc and super-Eddington regime. For $z > 6$ and for the luminosities that we are considering, the AGNs are in the thin disc or super-Eddington regime, so this SED is appropriate, although in future work, we plan to include a wider variety of SEDs for AGNs in different accretion regimes.

3.3 AGN obscuration and visible fractions

AGNs are understood to be surrounded by a dusty torus, which causes some of the radiation to be absorbed along some sightlines, and re-emitted at longer wavelengths. For simplicity, we assume that at a given wavelength, AGNs are either completely obscured or completely unobscured. The effect of obscuration can therefore be expressed as a visible fraction, which is the fraction of objects that are unobscured in a certain waveband at a given luminosity and redshift.

The fraction of obscured objects in the hard X-ray band is thought to be small, so for this work we assume that there is no obscuration at hard X-ray wavelengths. There is a population of so-called Compton-thick AGNs for which the column density of neutral hydrogen exceeds $N_{\rm H} \approx 1.5 \times 10^{24}$ cm^{-2}, which is the unit optical depth corresponding to the Thomson cross-section. Such objects are difficult to detect, even at hard X-ray wavelengths. The number of such objects is thought to be small, so we ignore their contribution for this work.

We calculate the visible fractions in the soft X-ray and optical bands using one of three observationally determined empirical relations from the literature, and also two more introduced in this work.

(i) The visible fraction of Hasinger (2008) is

$$f_{\rm vis} = 1 + 0.281\left[\log_{10}\left(\frac{L_{\rm HX}}{10^{43.75}\,{\rm ergs}^{-1}}\right)\right] - A(z), \quad (38)$$

where

$$A(z) = 0.279(1+z)^{0.62}. \quad (39)$$

[8]Note that equations (33) and (34) are corrected from Fanidakis et al. (2012, equation 10).

[9]Note that equation (36) is different to Fanidakis et al. (2012, equation 13).

$L_{\rm HX}$ is the hard X-ray luminosity in the observer frame and z is the redshift.[10] The redshift dependence of the visible fraction in this model saturates at $z \geq 2.06$ and the visible fraction is not allowed to have values below 0 or above 1. Because the observational data on which this obscuration model is based only extend to $z = 2$, we extrapolate the model to $z > 2$ using $L_{\rm HX}$ as the rest-frame hard X-ray band at $z = 2$, i.e. 6–30 keV. For this obscuration model, if an object is obscured at soft X-ray wavelengths, then it is also assumed to be obscured at optical/UV wavelengths.

(ii) Hopkins, Richards & Hernquist (2007) derive a visible fraction of the form:

$$f_{\rm vis} = f_{46}\left(\frac{L_{\rm bol}}{10^{46}{\rm erg\,s^{-1}}}\right)^{\beta}, \tag{40}$$

where f_{46} and β are constants for each band. For the B band, $[f_{46}, \beta]$ are [0.260, 0.082], and for the soft X-ray band, $[f_{46}, \beta]$ are [0.609, 0.063]. This model does not require a high-redshift extrapolation, as it depends only on bolometric luminosity.

(iii) Aird et al. (2015) observationally determine a visible fraction for soft X-rays of the form:

$$f_{\rm vis} = \frac{\phi_{\rm unabs}}{\phi_{\rm unabs} + \phi_{\rm abs}}, \tag{41}$$

where $\phi_{\rm unabs}$, the number density of unabsorbed sources, and $\phi_{\rm abs}$, the number density of absorbed sources, are given by

$$\phi = \frac{K}{\left(\frac{L_{\rm HX}}{L_\star}\right)^{\gamma_1} + \left(\frac{L_{\rm HX}}{L_\star}\right)^{\gamma_2}}, \tag{42}$$

where the constants for both cases are given in Table 3. As for the Hasinger (2008) obscuration model, if the object is obscured at soft X-ray wavelengths, then we assume that it is also obscured at optical/UV wavelengths. For this obscuration model, we extrapolate to high redshift such that for $z > 3$, the $L_{\rm HX}$ hard X-ray band is the rest-frame band for $z = 3$.

(iv) We also use visible fractions that are modified versions of Hopkins et al. (2007). These visible fractions also depend solely on $L_{\rm bol}$, but with different coefficients. These coefficients were derived by constructing a bolometric luminosity function from the luminosity functions at optical, UV, and X-ray wavelengths. We used the Marconi et al. (2004) bolometric corrections and selected coefficients for the visible fraction so as to create a resultant bolometric luminosity function with the scatter between points minimized. This is described in Appendix C. The first of these new obscuration relations, the 'low-z modified Hopkins', (LZMH) visible fraction for rest frame 1500 Å has the form:

$$f_{\rm vis,LZMH} = 0.15\left(\frac{L_{\rm bol}}{10^{46}{\rm erg\,s^{-1}}}\right)^{-0.1}, \tag{43}$$

and for the soft X-ray band it has the form:

$$f_{\rm SX,LZMH} = 0.4\left(\frac{L_{\rm bol}}{10^{46}{\rm erg\,s^{-1}}}\right)^{0.1}. \tag{44}$$

(v) The second of these modified Hopkins visible fractions, the '$z = 6$ modified Hopkins' (Z6MH) visible fraction was derived by fitting the GALFORM $z = 6$ luminosity functions at 1500 Å and in the soft X-ray band to the observational estimates. This visible fraction is

$$f_{\rm vis,Z6MH} = 0.04, \tag{45}$$

for both rest-frame 1500 Å and soft X-rays.

[10]This empirical model and others we use from observational studies were derived using a slightly different cosmology from the one used in the P-Millennium, for simplicity we ignore the effect of this here.

3.4 Calculating model AGN luminosity functions

Typically, when one constructs a luminosity function from a simulation, only the AGNs that are switched on at each snapshot are included. However, if one does this, rarer objects with higher luminosities but which are only active for a short time are not sampled well. To probe the luminosity function for such objects, we average over a time window, $\Delta t_{\rm window}$. The time window should not be too large, as then we may miss the effect of multiple starbursts within the time window, because the simulation only outputs information on the most recent starburst. We select a time window for which the luminosity function using the time average method is converged to the luminosity function using only the AGN switched on at the snapshots. For the predictions here we set $\Delta t_{\rm window} = t_{\rm snapshot}/10$, where $t_{\rm snapshot}$ is the age of the universe at that redshift.

Each object is assigned a weight, w, given by

$$w = t_Q/\Delta t_{\rm window}, \tag{46}$$

where $t_Q = f_{\rm q} t_{\rm bulge}$ is the lifetime of the most recent quasar episode occurring within the time interval $\Delta t_{\rm window}$ as in Section 3.1. This weight is then applied to the number densities counting all AGNs occurring within the time interval $\Delta t_{\rm window}$, which then allows us to include higher luminosity events at lower number densities in the luminosity function. We show the effect of changing the value of $\Delta t_{\rm window}$, as well as the effect of simply using snapshot quantities on the predicted luminosity functions in Appendix D.

4 SMBH MASSES, ACCRETION RATES, AND SPINS

We start by showing some basic predictions from the new model for SMBH masses, accretion rates, and spins.

4.1 Black hole masses

In the left-hand panel of Fig. 3, we show the black hole mass function at $z = 0$ predicted by our model compared to observational estimates. The observations use indirect methods to estimate the black hole mass function, because of the lack of a large sample of galaxies with dynamically measured black hole masses. In Marconi et al. (2004) and Shankar et al. (2004, 2009), galaxy luminosity/velocity dispersion functions are combined with relations between black hole mass and host galaxy properties to estimate black hole mass functions. The predictions of the model fit well to the observational estimates within the observational errors, especially given that there will also be uncertainties on the black hole mass measurements and given the discrepancies between the observational estimates. The former means the predictions could still be consistent with observations at the high-mass end ($M_{\rm BH} \geq 10^9\,{\rm M}_\odot$).

The evolution of the black hole mass function for $0 < z < 12$ is shown in the right-hand panel of Fig. 3. Most of the SMBH mass is formed by $z \sim 2$, as the mass density of black holes is dominated by objects around the knee of the black hole mass function, and this knee is in place by $z \sim 2$. The dominant fuelling mechanism

Table 3. The parameters that correspond to the best-fitting visible fraction from Aird et al. (2015) where $\zeta = \log(1 + z)$. These parameter values have been obtained by private communication. See equations (41) and (42).

	Absorbed	Unabsorbed
$\log(\mathrm{K/Mpc^{-3}})$	$-4.48 + 3.38\zeta - 7.29\zeta^2$	$-5.21 + 3.21\zeta - 5.17\zeta^2$
$\log(L_\star/\mathrm{erg\,s^{-1}})$	$43.06 + 3.24\zeta - 1.59\zeta^2 + 0.43\zeta^3$	$43.80 - 0.57\zeta + 9.70\zeta^2 - 11.23\zeta^3$
$\log\gamma_1$	$-0.28 - 0.67\zeta$	$-0.44 - 1.25\zeta$
γ_2	2.33	2.32
β_{CT}	0.34	0.34

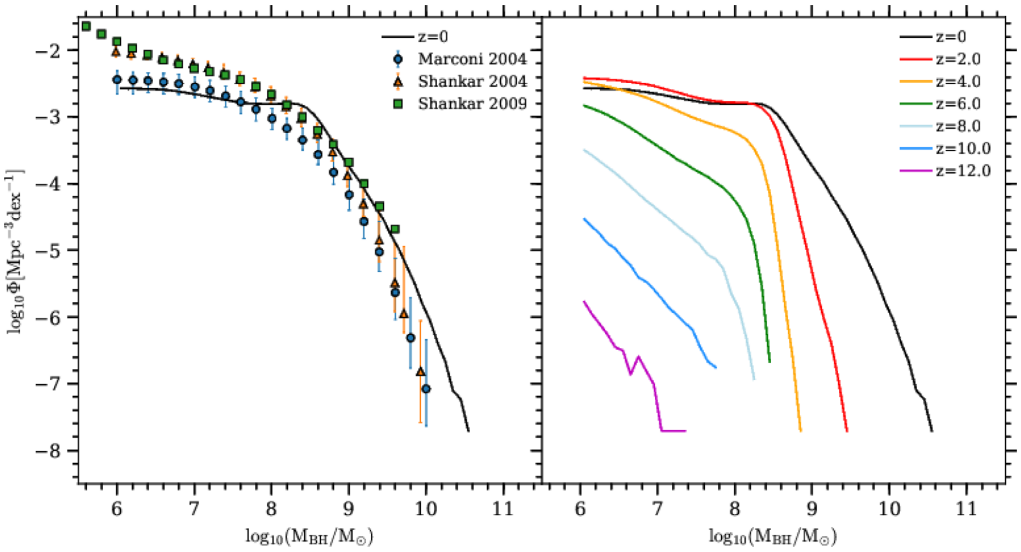

Figure 3. The black hole mass function. *Left-hand panel*: The predicted black hole mass function at $z = 0$ compared to observational estimates by Marconi et al. (2004); Shankar et al. (2004); and Shankar, Weinberg & Miralda-Escudé (2009). *Right-hand panel*: The evolution of the black hole mass function over the range $0 < z < 12$.

for growing the black hole mass density across all redshifts is gas accretion in starbursts triggered by disc instabilities, and disc instabilities play an important role in shaping the black hole mass function for $M_{\mathrm{BH}} < 10^8\,\mathrm{M}_\odot$. However, SMBH mergers are more important for determining the shape of the black hole mass function for $M_{\mathrm{BH}} > 10^8\,\mathrm{M}_\odot$, as they are the mechanism by which the largest SMBHs are formed. AGN feedback also plays an important role in shaping the black hole mass function at this high-mass end, by suppressing gas cooling and so slowing down the rate at which the SMBHs grow by cold gas accretion.

In Fig. 4, we show the 'active' black hole function at $z = 0$ compared to observational estimates from Schulze & Wisotzki (2010). In this observational estimate, active SMBHs are defined as AGNs radiating above a certain Eddington ratio ($L_{\mathrm{bol}}/L_{\mathrm{Edd}} > 0.01$). The flux limit in the observations results in the observational sample being incomplete for $M_{B_J} > -19$. The observational sample also only includes type-1 (unobscured) AGN. Therefore, we apply these selections to the model predictions, using the LZMH visible fraction to compare with this observational estimate of the active black hole mass function. We also present predictions where the selection on M_{B_J} has not been applied. The effect of the selection on M_{B_J} can be seen at the low-mass end ($M_{\mathrm{BH}} < 10^8\,\mathrm{M}_\odot$), where the dashed and solid lines diverge. While the model is in reasonable agreement with the observations at $M_{\mathrm{BH}} \sim 10^{8.5}\,\mathrm{M}_\odot$, the model generally underpredicts the active black hole mass function, although the model does reproduce the overall shape of the shape of the observational active black hole mass function. We found similar

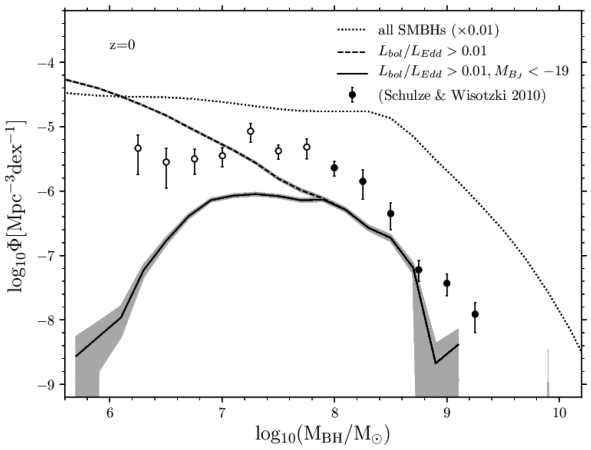

Figure 4. The active black hole mass function (solid line) at $z = 0$, compared to observational estimates from Schulze & Wisotzki (2010). We show predictions where active SMBHs are defined as AGNs brighter than a threshold Eddington ratio ($L_{\mathrm{bol}}/L_{\mathrm{Edd}} > 0.01$), using the LZMH visible fraction (cf. Section 3.3) (dashed line), and predictions also brighter than a threshold AGN absolute magnitude ($M_{\mathrm{B_J}} < -19$) (solid line). This is for appropriate comparison with the active black hole mass function in Schulze & Wisotzki (2010), where the open circles are the data points that suffer from incompleteness, while the filled circles are the data points that do not. We also show the total black hole mass function (dotted line) with the number density divided by 100, for comparison.

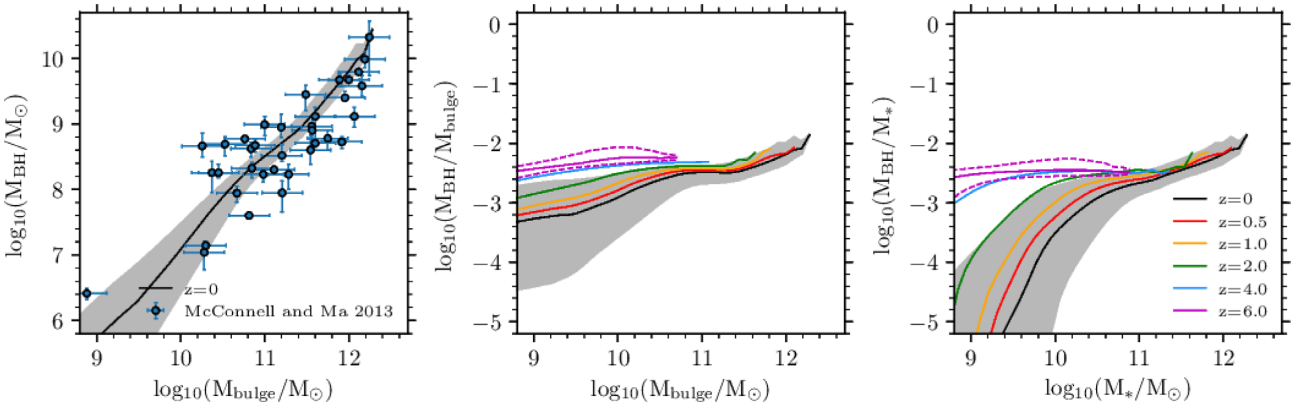

Figure 5. *Left-hand panel*: The predicted SMBH mass versus bulge stellar mass relation at $z = 0$ compared to observational data from McConnell & Ma (2013). The line represents the median of the predicted SMBH mass in bins of bulge mass and the shading denotes the 10–90 percentiles of the predicted distribution. *Middle panel*: The evolution of the median of the ratio of SMBH mass to bulge mass versus bulge mass relation with redshift for $z = 0, 0.5, 1, 2, 4, 6$. As in the left-hand panel, the grey shaded band is the 10–90 percentiles of the distribution for $z = 0$ and the purple dashed lines are the 10–90 percentiles of the distribution for $z = 6$. *Right-hand panel*: The evolution of the median of the ratio of SMBH mass to galaxy stellar mass versus galaxy stellar mass relation, with the lines representing the same redshifts as the middle panel as indicated by the legend.

results when comparing with other studies, such as those from SDSS (e.g. Vestergaard & Osmer 2009).

Fig. 5 shows the relation between SMBH mass and bulge or total stellar mass. In the left panel of Fig. 5 we show the predicted SMBH mass versus bulge mass relation compared to observational data from McConnell & Ma (2013). The predictions follow the observations well, with the scatter decreasing towards higher masses. BH–BH mergers contribute towards this decrease in scatter, as seen in Jahnke & Macciò (2011), although they are not the only contributing mechanism, with AGN feedback also affecting the scatter at the high-mass end.

In the middle panel of Fig. 5, we show the evolution of the ratio of SMBH mass to bulge mass ($M_{\rm BH}/M_{\rm bulge}$) versus bulge stellar mass for $0 < z < 6$, showing the scatter of the distribution for $z = 0$ and $z = 6$. As we go to higher redshift, the ratio $M_{\rm BH}/M_{\rm bulge}$ increases, as also seen in observations (e.g. Peng et al. 2006). The ratio $M_{\rm BH}/M_{\rm bulge}$ reflects the mechanism by which these two galaxy components form. At higher redshift, bulges grow mainly by starbursts, which also feeds the growth of SMBHs and so the distribution of the ratio $M_{\rm BH}/M_{\rm bulge}$ peaks at $f_{\rm BH}$ (the fraction of the mass of stars formed in a starburst accreted on to a black hole), with some scatter caused by mergers. At lower redshift the ratio $M_{\rm BH}/M_{\rm bulge}$ decreases, as galaxy mergers cause bulges to form from discs, but without growing the SMBHs. We also note how the scatter of the relation is lower at $z = 6$ than at $z = 0$ for all masses – by $z = 0$ galaxies have had more varied formation histories compared to the $z = 6$ population.

In the right-hand panel of Fig. 5, we show the evolution of the ratio of the SMBH mass to the galaxy stellar mass ($M_{\rm BH}/M_\star$) versus galaxy stellar mass for the redshift range $0 < z < 6$. Galaxies of larger stellar mass and the largest SMBHs form at late times, and at lower masses ($M_\star < 10^{11}\,{\rm M}_\odot$), $M_{\rm BH}/M_\star$ is smaller at later times. At lower masses, the ratio $M_{\rm BH}/M_\star$ decreases with time because the fraction of the stellar mass that is in the bulge decreases. This evolution slows down at $z < 1$. At higher masses ($M_\star > 10^{11}\,{\rm M}_\odot$), the stellar mass and SMBH mass stay on the same relation independent of redshift. It is in this regime that the AGN feedback is operational: in our model, we use the AGN feedback prescription of Bower et al. (2006) in which AGN feedback is only active where the hot gas halo is undergoing 'quasistatic' (slow) cooling. This has the effect that AGN feedback is only active for haloes of mass above $\sim 10^{12}\,{\rm M}_\odot$. The relation between SMBH mass and stellar mass at this high-mass end is caused by both AGN feedback and mergers, with neither mechanism dominant in establishing this relation.

4.2 Black hole accretion rates

In Fig. 6, we show the black hole mass accretion rate distribution, showing its evolution with redshift and split by fuelling modes: the hot halo mode, starbursts triggered by mergers and starbursts triggered by disc instabilities (see Section 2). The hot halo mode becomes more dominant at later times, because the hot halo mode requires long cooling times, and hence it occurs for massive haloes, and because dark matter haloes grow hierarchically, these large haloes only form at later times. The contribution from starbursts triggered by galaxy mergers peaks at $z \approx 2$. Starbursts triggered by mergers peak at a low-mass accretion rate, as seen in Fig. 6, albeit with a tail that extends to high $\dot{M}$. The peak at $\dot{M} \sim 10^{-6}\,{\rm M}_\odot\,{\rm yr}^{-1}$ is mostly due to minor mergers with mass ratios $0.05 < M_2/M_1 < 0.3$ (mergers with mass ratios in this range cause about three-quarters of the merger triggered starbursts at this mass accretion rate).[11] The contribution from starbursts triggered by disc instabilities increases as the redshift increases. Starbursts triggered by mergers typically have lower $\dot{M}$ values than starbursts triggered by disc instabilities. There are two reasons for this. First, the average stellar mass formed by bursts triggered by disc instabilities is higher than for bursts triggered by mergers, and this occurs because the average cold gas mass is higher for galaxies in which bursts triggered by disc instabilities occur. Secondly, the average bulge dynamical time-scale for starbursts triggered by disc instabilities is smaller than for those triggered by mergers due to the average bulge size being smaller for starbursts triggered by disc instabilities. The combination of these effects accounts for the lack of starbursts triggered by disc instabilities at the very lowest $\dot{M}$ values. The galaxies that host such starburst episodes would be below the mass at which the simulation is complete.

[11] Note that a mass ratio of 0.05 is assumed to be the lower threshold for starburst triggering in galaxy mergers (Lacey et al. 2016)

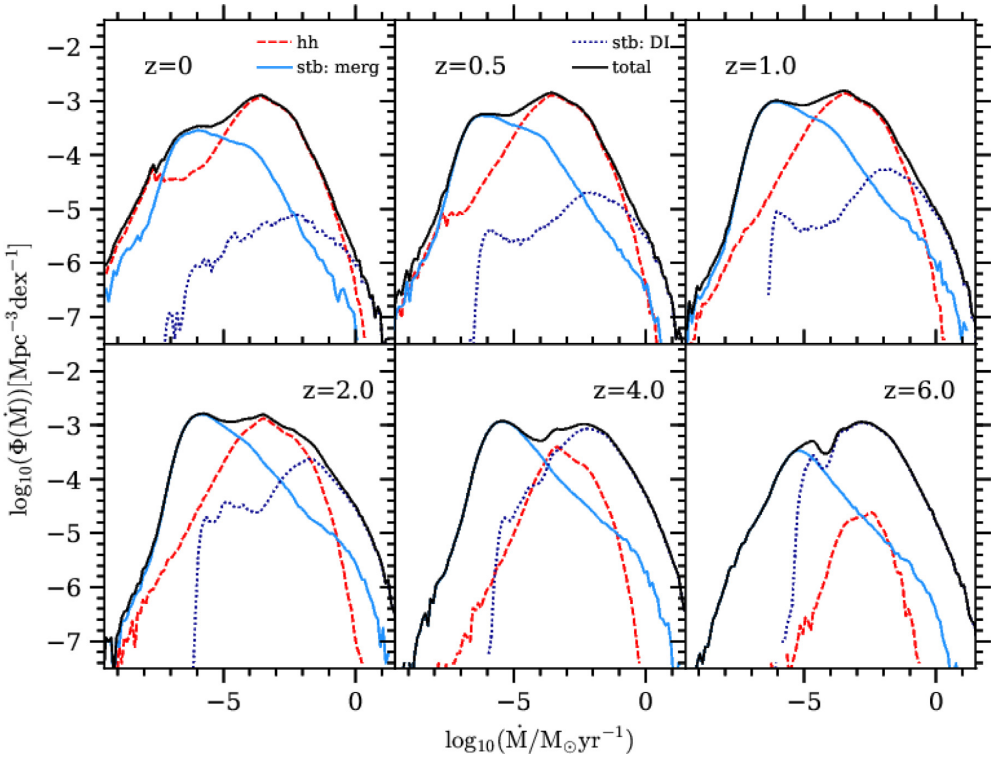

Figure 6. The distribution of black hole mass accretion rates for different redshifts (black solid line) split by contributions from hot halo mode (red dashed line), starbursts triggered by mergers (light blue solid line) and starbursts triggered by disc instabilities (dark blue dotted line). We have selected all black holes residing in galaxies of stellar mass, $M_\star > 10^6\,\mathrm{M}_\odot$, which is above the completeness limit of the simulation.

In Fig. 7, we show the evolution of the distribution of Eddington normalized mass accretion rate $\dot{M}/\dot{M}_{\rm Edd}$. We also show the predictions in different stellar mass ranges. Looking at the total distribution ($M_\star > 10^7\,\mathrm{M}_\odot$), for increasing redshift, the distribution shifts to somewhat higher values. This is seen as the number of objects with $\log(\dot{M}/\dot{M}_{\rm Edd}) < -2$ decreasing with increasing redshift, a peak at $\log(\dot{M}/\dot{M}_{\rm Edd}) \sim -1$ building up with increasing redshift and the number of objects with $\log(\dot{M}/\dot{M}_{\rm Edd}) > 0$ increasing with increasing redshift. The different bins of stellar mass have different distributions of $\dot{M}/\dot{M}_{\rm Edd}$, and evolve differently. At $z = 0$, the lowest bin in stellar mass ($10^7\,\mathrm{M}_\odot < M_\star < 10^9\,\mathrm{M}_\odot$) shows a broad distribution around a peak at $\log(\dot{M}/\dot{M}_{\rm Edd}) \approx -1.5$, the middle bin in stellar mass ($10^9\,\mathrm{M}_\odot < M_\star < 10^{11}\,\mathrm{M}_\odot$) also shows a broad distribution, but with a peak at $\log(\dot{M}/\dot{M}_{\rm Edd}) \approx -3$ and also has features at $\log(\dot{M}/\dot{M}_{\rm Edd}) \approx -1.5$ and $\log(\dot{M}/\dot{M}_{\rm Edd}) \approx -0.5$. The distribution in the highest stellar mass bin ($M_\star > 10^{11}\,\mathrm{M}_\odot$) peaks at lower value of $\log(\dot{M}/\dot{M}_{\rm Edd}) \approx -4$, but has fewer objects at high Eddington ratios than the lower stellar mass bins. The distribution in the highest stellar mass bin peaks at a lower Eddington ratio because this is where the hot halo mode is operational, so SMBHs are typically quiescently accreting.

As redshift increases, the number density at the peak in the $\dot{M}/\dot{M}_{\rm Edd}$ distribution for the lowest stellar mass bin increases, such that by $z = 6$, the peak for the lowest stellar mass bin and the middle stellar mass bin are both at $\log(\dot{M}/\dot{M}_{\rm Edd}) \approx -0.5$. The number of objects in the highest stellar mass bin decreases strongly at high redshift, so the hot halo mode is much less prevalent at higher redshift, $z > 3$.

We also have compared the predicted Eddington luminosity ratio, ($L_{\rm bol}/L_{\rm Edd}$) distribution at $z = 6$, to the observational data compiled in Wu et al. (2015, Fig. 4). The $L_{\rm bol}/L_{\rm Edd}$ distribution at $z = 6$ from GALFORM has a median and 10–90 percentiles at $4.3^{+4.3}_{-3.0}$ for AGNs with $L_{\rm bol} > 10^{46}\,\mathrm{ergs}^{-1}$) and $8.6^{+3.5}_{-3.5}$ for AGNs with $L_{\rm bol} > 10^{47}\,\mathrm{ergs}^{-1}$, whereas the $L_{\rm bol}/L_{\rm Edd}$ median and 10–90 percentiles in Wu et al. (2015) is $1.0^{+1.8}_{-0.4}$ for a mixture of samples with $L_{\rm bol} > 10^{46}\mathrm{erg\ s}^{-1}$. The predicted $L_{\rm bol}/L_{\rm Edd}$ are somewhat larger than the observational estimate. One possible reason for the different distributions is systematic uncertainties in the black hole mass estimates in the observations. We plan to conduct a more detailed investigation in future work.

4.3 Black hole spins

In Fig. 8, we show the SMBH spin distribution predicted by the model for both the prolonged and chaotic accretion modes. Note that a here represents the magnitude of the spin. The low-mass end of the spin distribution [$6 < \log_{10}(M_{\rm BH}/\mathrm{M}_\odot) < 8$] is dominated by accretion spin-up, whereas the high-mass end ($8 < \log_{10}(M_{\rm BH}/\mathrm{M}_\odot) < 10$) is dominated by merger spin-up. For prolonged mode accretion, the coherent accretion spin-up means that SMBHs quickly reach their maximum spin value, giving rise to a population of maximally spinning SMBHs at low mass. At high masses, the average spin value is lower because of SMBH mergers. This is because even if two maximally spinning SMBHs merge, the result is typically an SMBH with a lower spin value because of misalignment between the black hole spins and the orbital angular momentum. For chaotic mode accretion, the accretion direction is constantly changing and so the accretion spin-up leads to SMBHs with lower median spin values ($a \approx 0.4$), compared to prolonged accretion. The spin values are not zero in the chaotic mode, as one may be tempted to expect, because the accretion spin-up is

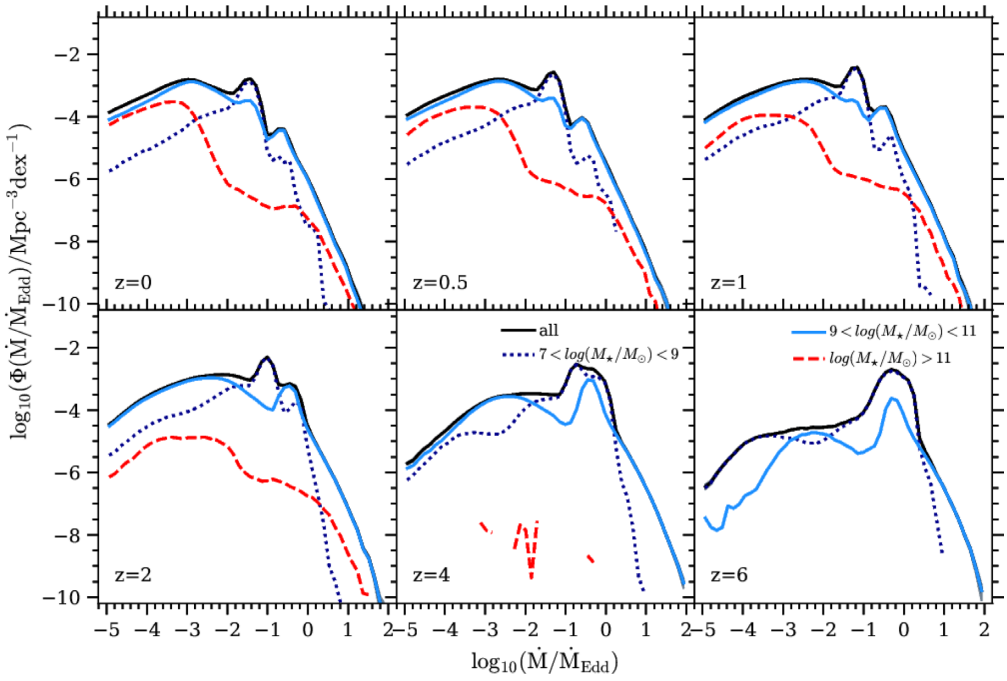

Figure 7. The distribution of Eddington ratio in terms of mass accretion rate, $\dot{M}/\dot{M}_{\rm Edd}$, evolving with redshift. Shown are all objects with stellar mass, $M_\star > 10^7$ M$_\odot$ (black solid line), objects with stellar mass 10^7 M$_\odot$ $< M_\star < 10^9$ M$_\odot$ (dark blue dotted line), objects with stellar mass 10^9 M$_\odot$ $< M_\star < 10^{11}$ M$_\odot$ (light blue solid line), and objects with stellar mass $M_\star > 10^{11}$ M$_\odot$ (red dashed line).

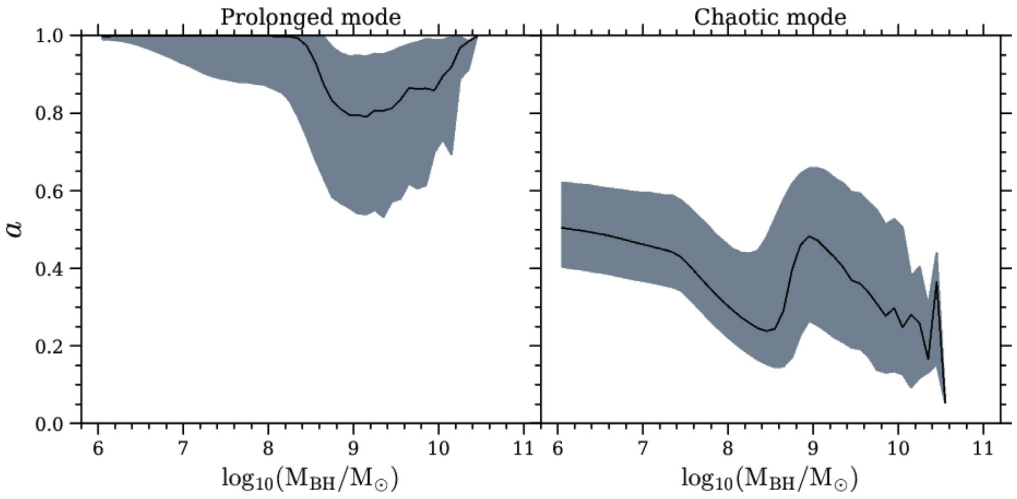

Figure 8. The predicted SMBH spin distributions at $z = 0$ for prolonged (left-hand panel) and chaotic (right-hand panel) accretion modes. The line represents the median value of the magnitude of the spin for that SMBH mass, and the shading represents the 10–90 percentile range of the distribution.

more efficient if the accretion disc and SMBH spin are in the same direction compared to the case of anti-alignment (King et al. 2008). The mean value of the SMBH spin decreases with increasing black hole mass at this low-mass end for chaotic mode accretion as also reported in King et al. (2008). At the high-mass end, the increase in average spin at $M_{\rm BH} \sim 10^9$ M$_\odot$ is due to spin-up by BH mergers. Two slowly spinning SMBHs typically form a higher spin SMBH when they merge due to the angular momentum of the orbit between them.

One of the conclusions of Fanidakis et al. (2011) was that for chaotic mode accretion, smaller SMBHs will have lower spin values ($\bar{a} \approx 0.15$), whereas larger SMBHs will have higher spin values ($\bar{a} \approx 0.7 - 0.8$). Our new analysis predicts that for chaotic mode accretion SMBHs will generally have moderate spin values, $\bar{a} \approx 0.4$, yielding radiative accretion efficiencies of $\epsilon \approx 0.075$, not too dissimilar from the value of $\epsilon \approx 0.1$ required by the Soltan (1982) argument. However, the average radiative accretion efficiency implied by prolonged mode accretion is $\epsilon \approx 0.4$ in tension with the Soltan (1982) argument.

The chaotic mode spin distribution is different to that in Fanidakis et al. (2011) because the equations for SMBH spin-up by gas accretion have changed from that paper (causing higher spin values at the low-SMBH mass end) and because the directions for the spin-up due to SMBH mergers are sampled from the surface of a sphere as opposed to the circumference of a circle, leading to lower spin values at the high SMBH mass end.

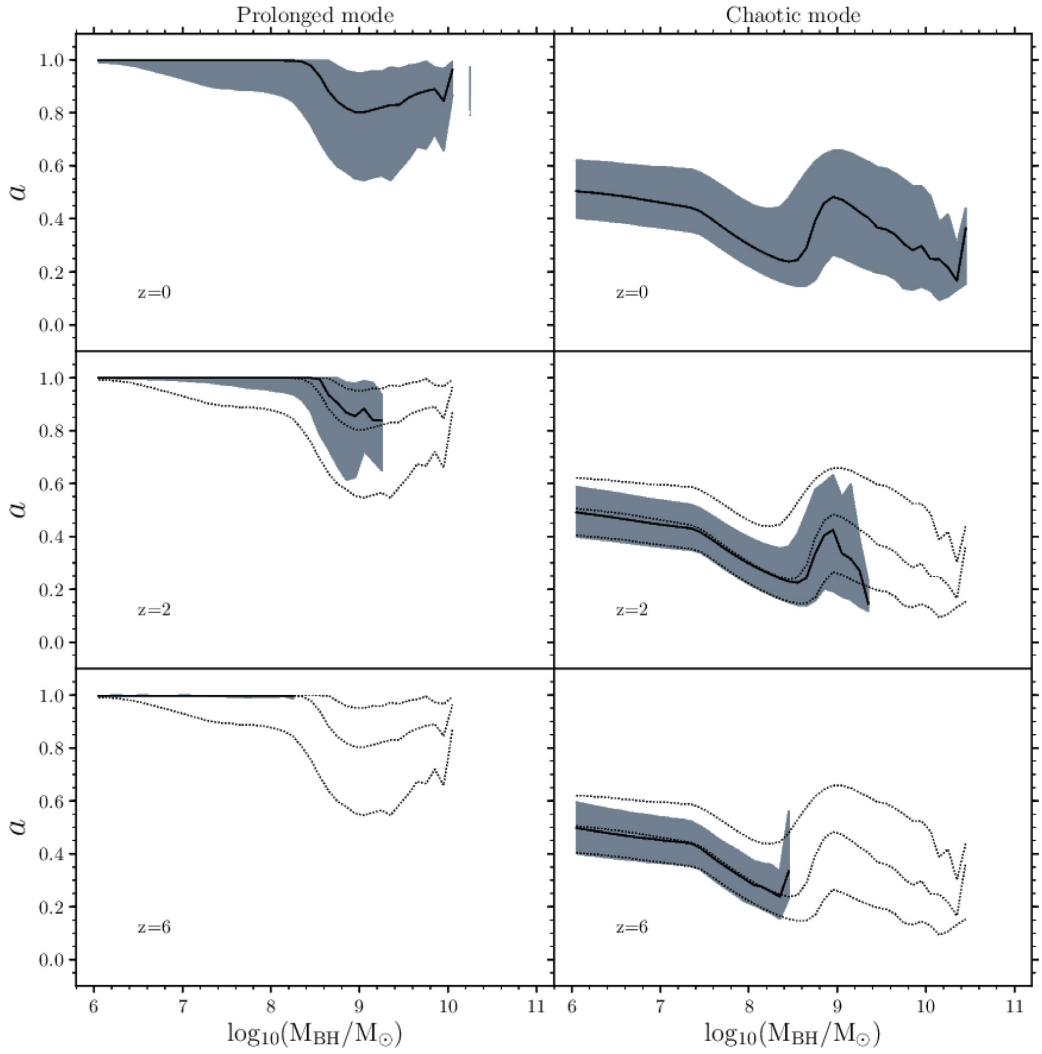

Figure 9. The predicted evolution of the SMBH spin distribution for prolonged mode (left-hand panels) and chaotic mode (right-hand panels). Results are shown for $z = 0$, 2, 6. The lines and shading have the same meaning as in the previous figure, with the dotted line representing the median and percentiles for that accretion mode at $z = 0$.

We then show the evolution of the SMBH spin distribution for the prolonged and chaotic modes in Fig. 9. The black hole spin versus black hole mass relation shows negligible evolution for both modes, with the median black hole spin at any black hole mass approximately the same over the range $z = 0$–6. For both modes, the scatter of the distribution decreases with increasing redshift, with the scatter for the prolonged mode decreasing much more than the scatter for the chaotic mode. For the prolonged mode, by $z = 6$, nearly all of the black holes with $M_{\rm BH} < 10^8\,{\rm M}_\odot$ have the maximal spin permitted by the model. Also, there is a lack of high mass, $M_{\rm BH} > 3 \times 10^8\,{\rm M}_\odot$, black holes at $z = 6$ for both modes. This is due to a low abundance of high-mass galaxies at $z = 6$.

We show how typical black holes evolve in the chaotic mode (the standard choice for this analysis) for four different black hole masses in Figs 10 and 11. When we generate each black hole history, we only follow the largest progenitor black hole back in time when two or more black holes merge. In the upper panel of Fig. 10, we show the evolution of the black hole mass through time evolution for these objects, where the time is measured from the big bang. Some of the features discussed for the black hole mass function in Fig. 3 can be seen here, such as how most of the SMBH mass is assembled at early times, and how the very largest black holes build up gradually at late times. It can also be seen how the larger SMBHs generally grow their mass quickest, with smaller SMBHs generally growing later. This is seen in Fig. 10 where the SMBH of mass $M_{\rm BH} = 5.47 \times 10^6\,{\rm M}_\odot$ reaches 40 per cent of its final mass at 9 Gyr, whereas the SMBH of mass $M_{\rm BH} = 8.43 \times 10^7\,{\rm M}_\odot$ reaches 60 per cent of its final mass at 6 Gyr, and the SMBH of mass $M_{\rm BH} = 4.13 \times 10^8\,{\rm M}_\odot$ reaches 80 per cent of its final mass at 2 Gyr. However, the SMBH of mass $M_{\rm BH} = 2.68 \times 10^9\,{\rm M}_\odot$ grows more gradually.

In the lower panel of Fig. 10, we show the evolution of SMBH spin through time. SMBHs of different masses generally show the same trends as their spin evolves through time. At early times, the black holes are smaller and so the spin values will change dramatically (with a changing between 0 and 0.8) if there is an accretion or merger event, whereas at later times, the spin values do not change as dramatically (a only varies by about 0.1 for each event) with time. The spin values generally converge on a moderate value ($a \approx 0.2$–0.6) at late times.

In Fig. 11, we show the evolution of the black holes through the spin versus mass plane. First, the black holes are spun up to high spins by mergers at small masses. Then the black holes of different

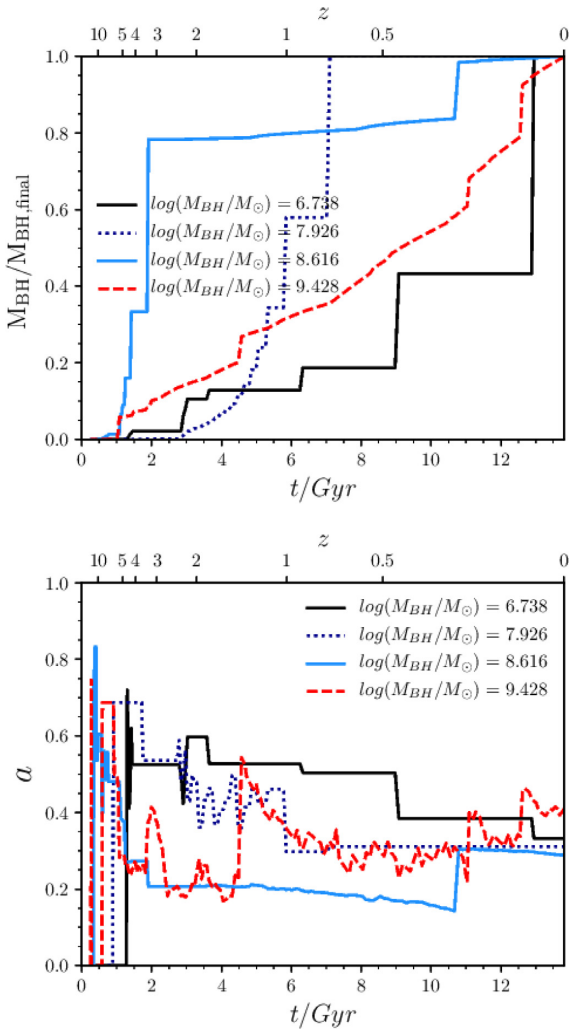

Figure 10. *Upper panel:* The evolution of the ratio of SMBH mass to the SMBH at $z = 0$ versus time. *Lower panel:* The evolution of SMBH spin versus time. In both panels, we show examples of SMBHs with $z = 0$ masses of $M_{BH} = 5.47 \times 10^6$ M$_\odot$ (black solid line), $M_{BH} = 8.43 \times 10^7$ M$_\odot$ (dark blue dotted line), $M_{BH} = 4.13 \times 10^8$ M$_\odot$ (light blue solid line), and $M_{BH} = 2.68 \times 10^9$ M$_\odot$ (red dashed line). The same objects are plotted in both panels.

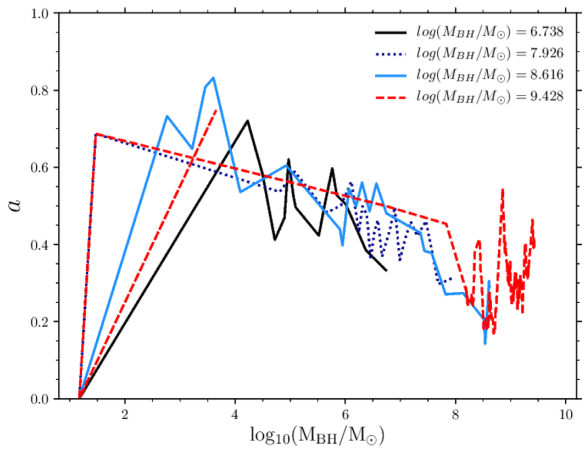

Figure 11. The evolution of four different mass SMBHs through the spin versus mass plane. The final SMBH masses at $z = 0$ are the same as plotted in Fig. 10.

masses generally show a similar evolution through the spin versus black hole mass plane as they evolve from high spins at lower black hole masses to lower spins at higher black hole masses, as they accrete gas by chaotic mode accretion. For the two largest black hole masses, there is an additional feature, as the black hole spin increases at the very highest masses. This is a result of the black holes merging with other black holes following their host galaxies merging.

4.4 AGN luminosities and black hole/galaxy properties

Before comparing the predicted AGN luminosity functions to observational estimates, we first show the dependence of AGN luminosities on some different galaxy properties.

First in the left-hand panel of Fig. 12, we show the dependence of bolometric luminosity on halo mass, where the points are coloured by the density of points. Each halo mass can host an AGN up to $L_{bol} \sim 10^{44}$ erg s^{-1}, with the brightest AGNs not residing in the largest haloes, but instead in haloes of mass $M_{halo} \sim 10^{12}$ M$_\odot$. This is a result of how in the model, AGN activity is inhibited in the largest haloes due to AGN feedback (cf. Fanidakis et al. 2013). The overall distribution is bimodal, which is a result of the two primary fuelling modes. The AGNs at $M_{halo} \lesssim 10^{12.5}$ M$_\odot$ are mostly fuelled by starbursts triggered by disc instabilities, whereas the AGNs at $M_{halo} \gtrsim 10^{12.5}$ M$_\odot$ are mostly fuelled by hot halo mode accretion. AGNs fuelled by starbursts triggered by mergers make a minor contribution to both parts of this distribution. Hot halo mode accretion fuels the objects at the peak of the 2D distribution in this plane seen at $M_{halo} \approx 10^{13}$ M$_\odot$ and $L_{bol} \approx 10^{42}$ erg s^{-1}. The peak of the distribution of objects fuelled by starbursts triggered by disc instabilities is at $M_{halo} \approx 10^{11.5}$ M$_\odot$ and $L_{bol} \approx 10^{43.5}$ erg s^{-1}, while the peak in the distribution for starbursts triggered by mergers is at $M_{halo} \approx 10^{11.5}$ M$_\odot$ and $L_{bol} \approx 10^{42}$ erg s^{-1}.

In the right-hand panel of Fig. 12, we show the dependence of bolometric luminosity on stellar mass. There is more of a correlation between bolometric luminosity and stellar mass than between bolometric luminosity and halo mass. The brightest AGNs in the model do not live in the largest stellar mass galaxies, but rather reside in galaxies of $M_\star \sim 10^{11}$ M$_\odot$. This distribution also shows a bimodality, where generally the objects at lower masses ($M_\star < 3 \times 10^{10}$ M$_\odot$) are fuelled by the starburst mode, while objects at higher masses ($M_\star > 3 \times 10^{10}$ M$_\odot$) are fuelled by the hot halo mode, although there is some overlap between the two. For the starburst mode, the peak of the distribution for starbursts triggered by disc instabilities and the peak of the distribution for starbursts triggered by mergers are both at stellar mass $M_\star \approx 3 \times 10^9$ M$_\odot$. This peak is at $L_{bol} \approx 10^{43}$ erg s^{-1} for disc instabilities, whereas for mergers this peak is at $L_{bol} \approx 10^{42}$ erg s^{-1}. Starbursts triggered by mergers do also occur for galaxies of stellar mass $M_\star > 10^{11}$ M$_\odot$, whereas starbursts triggered by disc instabilities do not occur for galaxies of this mass.

In Fig. 13, we show the dependence of AGN bolometric luminosity on the duration of the AGN episode. The distribution peaks at $t_{AGN} \approx 10^{7.5}$ yr and $L_{bol} \approx 10^{42}$ erg s^{-1}, with objects with luminosities $L_{bol} < 10^{44}$ erg s^{-1} having a wide range of durations of the AGN episodes. However, the brightest objects at $L_{bol} \approx 10^{48}$ erg s^{-1} all have durations of $t_{AGN} \approx 10^6$ yr with an anticorrelation between duration of the AGN episode and the AGN luminosity. This anticorrelation arises because, in general, shorter AGN episodes lead to higher AGN luminosities.

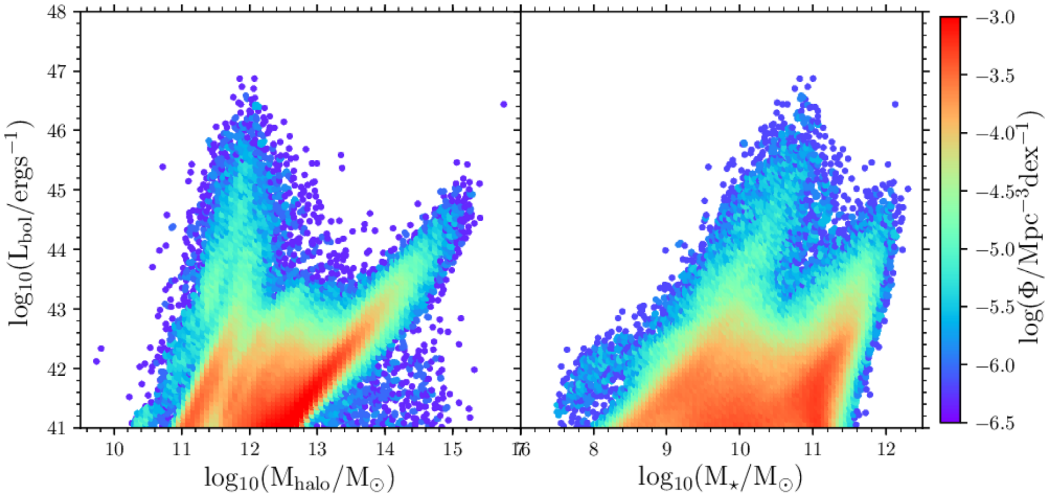

Figure 12. *Left-hand panel:* A scatter plot of AGN bolometric luminosity versus halo mass at $z = 0$. The points are coloured by the density of objects in this plane, where red indicates a high density of objects while blue indicates a low density of objects. *Right-hand panel:* As in the left-hand panel but showing bolometric luminosity versus stellar mass.

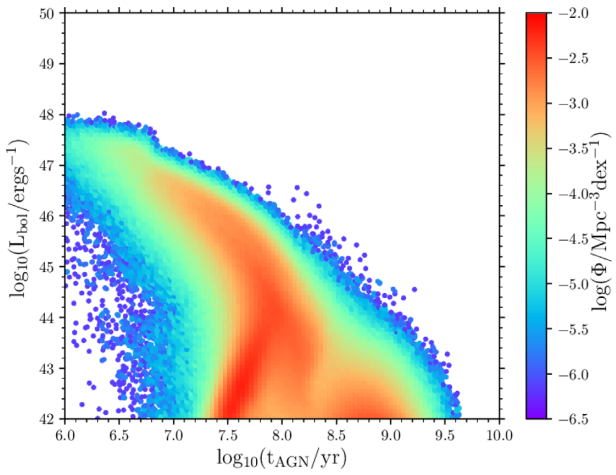

Figure 13. As in Fig. 12 but showing the dependence of AGN bolometric luminosity on the duration of the AGN episode for starburst mode fuelled AGN only.

5 EVOLUTION OF THE AGN LUMINOSITY FUNCTION AT $z < 6$

We first discuss the evolution of the predicted AGN luminosity function, as it is the simplest to predict, and then the AGN luminosity functions at different wavelengths, which depend on bolometric and obscuration corrections.

5.1 Bolometric luminosity function

We present the predicted bolometric luminosity function compared to our observationally estimated bolometric luminosity function constructed from multiwavelength data. This observationally estimated bolometric luminosity function is described in Appendix C, and is compared to other observational estimates in Appendix C.

The model for SMBH evolution and AGN luminosity also involves some free parameters additional to those in the galaxy formation model, as shown in Table 2. We have calibrated the values of $f_{\rm q}$ and $\eta_{\rm Edd}$, and found that the best-fitting values are those adopted in Fanidakis et al. (2012). We show the effect of varying these parameters in Figs E1 and E2. We also slightly adjust the values of $\alpha_{\rm ADAF}$ and $\alpha_{\rm TD}$ from 0.087 to 0.1. This is for simplicity and to keep the values in line with MHD simulations (e.g. Penna et al. 2013). The value of $\delta_{\rm ADAF}$ has been updated from Fanidakis et al. (2012, cf. Section 3.1)

In Fig. 14, the predictions (where the black line is the sum of the contributions from all accretion modes) compare well to the observational bolometric luminosity function across the range of redshifts and for the luminosities shown. Exceptions include the faint end at high redshift where the model overpredicts the observations by 0.5 dex for $L_{\rm bol} < 10^{46}\,{\rm erg\,s^{-1}}$ for $z > 4$, and the faint end at low redshift where the model underpredicts the observations for $L_{\rm bol} < 10^{45}\,{\rm erg\,s^{-1}}$ and $z < 0.5$ by 0.5 dex. The underpredictions at the faint end at low redshift may be because the ADAF radiative accretion efficiency is lower than the thin disc accretion efficiency, leading to lower luminosities (see Fig. E5 for a prediction using only a thin disc accretion efficiency for all values of $\dot{m}$). Alternatively, this discrepancy might be resolved by assuming an accretion time-scale with a dependence on accreted gas mass or black hole mass. For a different model, Shirakata et al. (2019) obtain a better fit to the hard X-ray luminosity function at low luminosity and low redshift by doing this. In general, our model is a good match to these observations across a broad range.

We also show in Fig. 14 the separate contributions to the AGN luminosity function from ADAFs ($\dot{m} < \dot{m}_{\rm crit,ADAF}$), thin discs ($\dot{m}_{\rm crit,ADAF} < \dot{m} < \eta_{\rm Edd}$), and super-Eddington objects ($\dot{m} > \eta_{\rm Edd}$). At low redshift, ADAFs dominate the faint end ($L_{\rm bol} < 10^{44}\,{\rm ergs^{-1}}$), thin discs dominate at intermediate luminosities ($10^{44}\,{\rm ergs^{-1}} < L_{\rm bol} < 10^{46}\,{\rm ergs^{-1}}$), and super-Eddington objects dominate the bright end ($L_{\rm bol} > 10^{46}\,{\rm ergs^{-1}}$). As we go to higher redshift, the ADAFs contribution to the luminosity function decreases: for $0 < z < 2$ the evolution is not that strong, although the contribution from ADAFs at each luminosity decreases slightly as we increase z in this range, whereas for $z > 2$, the evolution in the ADAF population is pronounced, and the number of ADAFs drops off sharply with increasing redshift. In contrast, the contribution from the thin disc population increases until $z \approx 2$, after which it remains approximately constant. At $z < 2$, there are not very many super-Eddington objects and so they make a fairly small contribution to

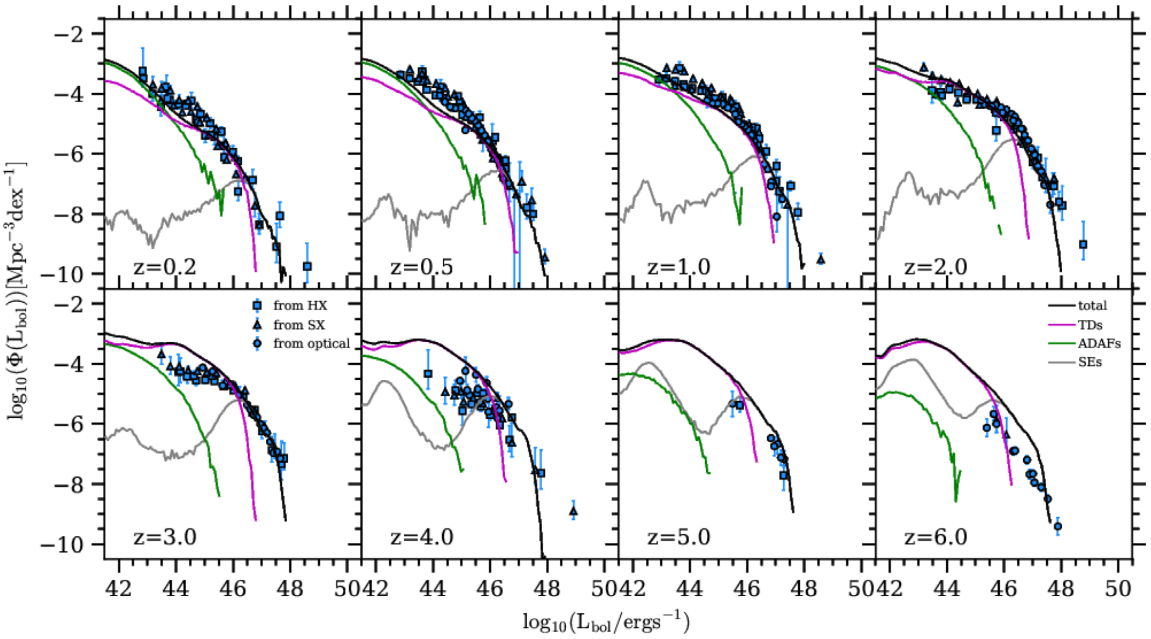

Figure 14. The AGN bolometric luminosity function predicted by our model (black line, with grey shading showing the Poisson errorbars) compared to our bolometric luminosity function constructed from the observations. We show the observational data indicating the wavelength of the data that was used to construct that particular point (squares – hard X-ray, triangles – soft X-ray, circles – optical). We split the total bolometric luminosity function by accretion mode into ADAFs (green), thin discs (purple), and super-Eddington objects (grey).

the luminosity function but their contribution increases at $z > 2$. The distribution of super-Eddington objects is bimodal, and for $z < 4$, the higher luminosity peak has a higher number density, while for $z > 4$, the lower luminosity peak has a higher number density. The bimodality is not due to the bimodality in the fuelling modes, as all the super-Eddington objects are fuelled by starbursts triggered by disc instabilities, but it seems to be caused by a bimodality in the bulge stellar mass. We plan to explore this issue in more detail in future work.

In Fig. 15, we split the AGN luminosity function by contributions from the hot halo mode, starbursts triggered by mergers and starbursts triggered by disc instabilities. At low redshift ($z < 2$), the faint end is dominated by the hot halo mode, whereas the bright end is dominated by starbursts triggered by disc instabilities. Starbursts triggered by mergers make a small contribution to the AGN bolometric luminosity function at low redshift. Starbursts triggered by disc instabilities typically have higher values of $\dot{M}$ and so higher luminosities compared to starbursts triggered by mergers, which is why they dominate the bright end.

The hot halo mode only operates in the most massive haloes, and so it only begins to significantly contribute to the AGN luminosity function for $z < 3$. The hot halo mode does not strongly evolve for $0 < z < 2$. For $z > 2$, starbursts triggered by disc instabilities dominate the AGN luminosity function, with starbursts from mergers not significantly contributing. This implies that the inclusion of black hole growth via disc instabilities is significant for reproducing AGN luminosity functions at high redshift.

A key aspect of the success of the GALFORM AGN model is the different channels of black hole growth, particularly the inclusion of disc instability triggered starbursts, that allow a good match to the AGN luminosity functions to be obtained. Other semi-analytic models do not necessarily include disc instabilities, which may explain why they do not reproduce AGN properties particularly well at high redshift (e.g. Bonoli et al. 2009; Menci et al. 2013; Neistein & Netzer 2014; Enoki et al. 2014). The effect of disc instabilities on the AGN predictions at $0 < z < 6$ is shown in Fig. E3 and the effect on galaxy properties is shown in Lacey et al. (2016).

We show the effect on the AGN bolometric luminosity function of changing between chaotic mode (our standard choice) and prolonged mode in Fig. 16. In the prolonged mode, SMBH spins are generally higher (see Fig. 8), which results in a higher radiative accretion efficiency leading to higher bolometric luminosities.[12]

5.2 Luminosity functions at different wavelengths

We use the SED template described in Section 3.2 and visible fractions described in Section 3.3 to make predictions for the luminosity function in the rest-frame hard X-ray, soft X-ray, and 1500 Å bands. In Fig. 17, we compare our hard X-ray predictions to observational data. The model is generally in good agreement with the observational data, particularly in the range $1 < z < 3$. For $L_{HX} < 10^{44}\,erg\,s^{-1}$ at $z < 0.5$, the model underpredicts the observations by about 0.5 dex, and for $L_{HX} < 10^{44}\,erg\,s^{-1}$ at $z > 3$, the model overpredicts the observations by about 1 dex. The former discrepancy corresponds to the model bolometric luminosity function underpredicting the observations in the same redshift and luminosity regime, and the latter also corresponds to the bolometric luminosity function slightly overpredicting the observational estimates in that regime, but may also be influenced by our assumption that there is no obscuration for hard X-ray sources. This assumption may be not valid for the high-redshift Universe; more observations are needed to constrain the obscuration effect on hard X-rays.

[12]Note that the shape of the luminosity function changes little between the two models.

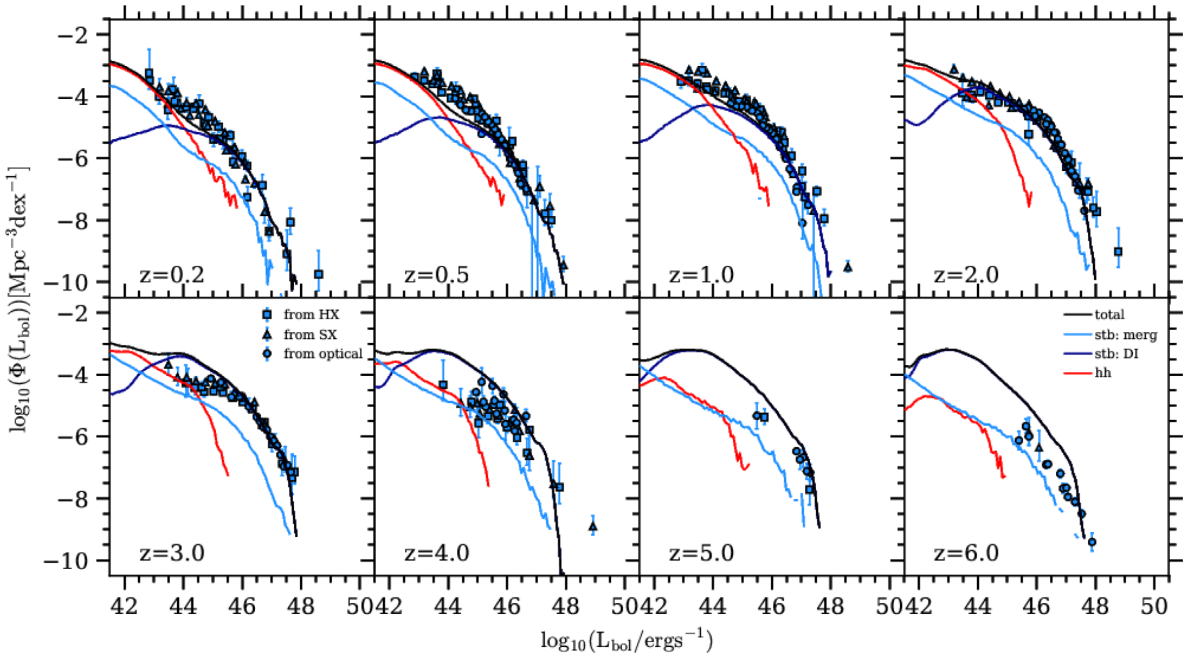

Figure 15. The AGN bolometric luminosity function as Fig. 14, but split by the fuelling mode: starbursts triggered by mergers (light blue), starbursts triggered by disc instabilities (dark blue), hot halo mode (red).

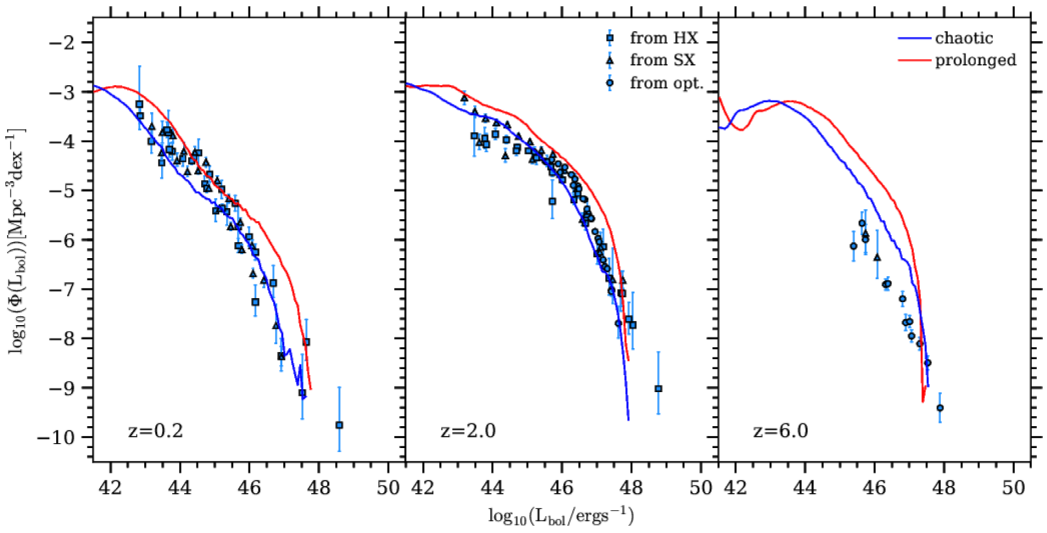

Figure 16. The effect of changing between chaotic (blue) and prolonged (red) mode on the AGN bolometric luminosity function at $z = 0.2$, 2, 6.

Our soft X-ray predictions are compared to observations in Fig. 18. The predicted luminosity function without taking into account obscuration is shown alongside the model with the visible fractions of Hopkins et al. (2007), Hasinger (2008), Aird et al. (2015), and our observationally determined LZMH model. The luminosity functions with different visible fractions are very similar except for $L_{SX} < 10^{44}$ erg s^{-1}. The LZMH model fits best to the observations in the range $1 < z < 2$. At higher redshifts and lower luminosities, the visible fraction in the Hasinger (2008) model drops to zero, which causes the corresponding drop off in the luminosity function for that obscuration model.

Our 1500 Å predictions are shown in Fig. 19 compared to observational estimates. These have been converted to 1500 Å – the conversions made are detailed in Appendix B. There is a strong dependence of the predictions on the assumed obscuration model. Our predictions are a good fit to observations at $z \approx 2$ if we adopt the Hasinger (2008) visible fraction, whereas our observationally determined LZMH model fits best for $z \approx 4$. The reason for this difference is likely to be because Hasinger (2008) fitted their obscuration model at lower redshift, whereas we are trying to fit for $z = 0$–6 with our LZMH visible fraction. Therefore, unsurprisingly, the different visible fractions are likely to fit better in different redshift ranges.

We present the soft X-ray and optical luminosity functions at $z = 6$ in Fig. 20. The predicted soft X-ray luminosity function exceeds the observations at $z = 6$ as a result of the model

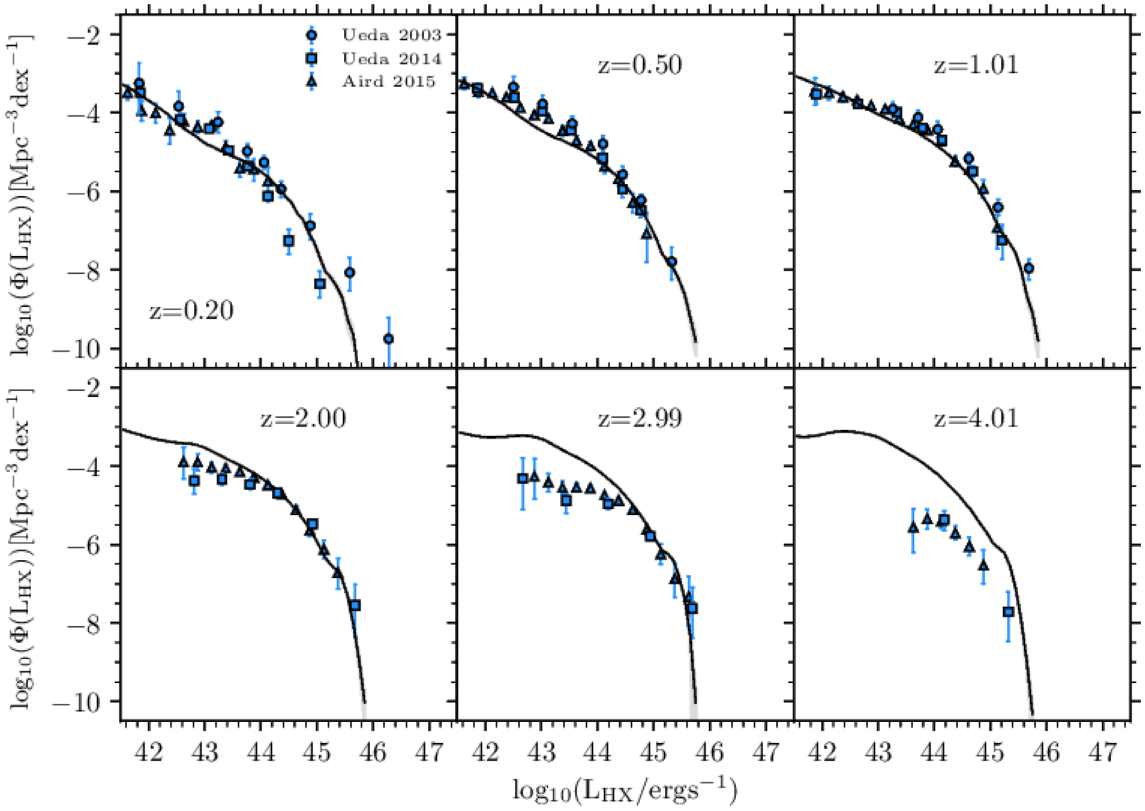

Figure 17. The rest-frame hard X-ray luminosity function predicted by the model (black line) compared to observational studies from Ueda et al. (2003) (circles), Ueda et al. (2014) (squares), and Aird et al. (2015) (triangles).

bolometric luminosity function overpredicting the observations. For the optical luminosity function, while the model gives an acceptable fit to observations of the optical luminosity function at $z = 4$, it overpredicts the number of AGNs compared to the observed luminosity function at $z = 6$. This is a result of the model not strongly evolving in the redshift interval $z = 4$–6, while the observations indicate a stronger evolution in this redshift interval (Jiang et al. 2016). These discrepancies could be due to a variety of reasons. We suggest two possible explanations for this discrepancy and two corresponding variants on the model that provides a better fit to the observations at $z = 6$.

First, the discrepancy could be due to the obscuration model. At $z = 6$, the visible fraction is not constrained by any observations, and so in Fig. 20, we present predictions with a lower visible fraction at $z = 6$, which give a better fit to the $z = 6$ optical luminosity function. We show predictions for the standard model with two obscuration models: the LZMH visible fraction and the Z6MH visible fraction (cf. Section 3.3). The Z6MH visible fraction needed to fit $z = 6$ is about a quarter of the LZMH visible fraction at $z < 6$. Thus, $z > 6$ QSOs could be much more obscured than $z < 6$ QSOs.

Secondly, the discrepancy could be due to black hole accretion being less efficient at high redshift. While the model for black hole accretion has been calibrated at low redshift, the conditions for black hole accretion could be different at higher redshift. We therefore present a model with parameters that have been modified compared to the original calibration on observed data at low redshift. We change the parameter $f_{\rm BH}$, which sets the fraction of mass accreted on to a black hole in a starburst event and the parameter $\eta_{\rm Edd}$, which controls the degree of super-Eddington luminosity suppression. In the fiducial model, $f_{\rm BH} = 0.005$ and $\eta_{\rm Edd} = 4$. $f_{\rm BH} = 0.002$ and $\eta_{\rm Edd} = 16$ give a better fit to the observations of the 1500 Å luminosity function at $z = 6$ in Fig. 20. However, we note that $\eta_{\rm Edd} = 16$ means that there is very little super-Eddington luminosity suppression, whereas the 'slim disc' model for super-Eddington sources predicts significant super-Eddington luminosity suppression. We refer to this model as the 'low accretion efficiency model'. In this model, we use the LZMH visible fraction.

Both of these alternative models are in better agreement with observations of the 1500 Å AGN luminosity function at $z = 6$ than our standard model, and so we will use them for future studies investigating AGNs observed in future surveys.

5.3 Comparison with hydrodynamical simulations

An alternative theoretical approach for simulating galaxy formation is hydrodynamical simulations. A few of these simulations have been used to make predictions for the evolution of AGN luminosity functions through time. We give a brief comparison to some of these here.

The bolometric luminosity function predicted by the model in Hirschmann et al. (2014) over the redshift range $0 < z < 5$ is shown in their fig. 8. When compared to Hopkins et al. (2007), their model is a good fit to the observations at $z = 0.1$, but overpredicts the observations at the faint end at $z = 2$, and underpredicts the observations at $z = 5$. When comparing their results to the model presented in this paper (cf. Fig. 14), our model agrees similarly well with the observations for $z < 2$, and with better agreement to the observations for $z > 2$. For example, at $z = 4$, at $L_{\rm bol} = 10^{46}$ ergs^{-1} (around the knee of the luminosity function at this redshift), our model agrees within 0.5 dex with the observed

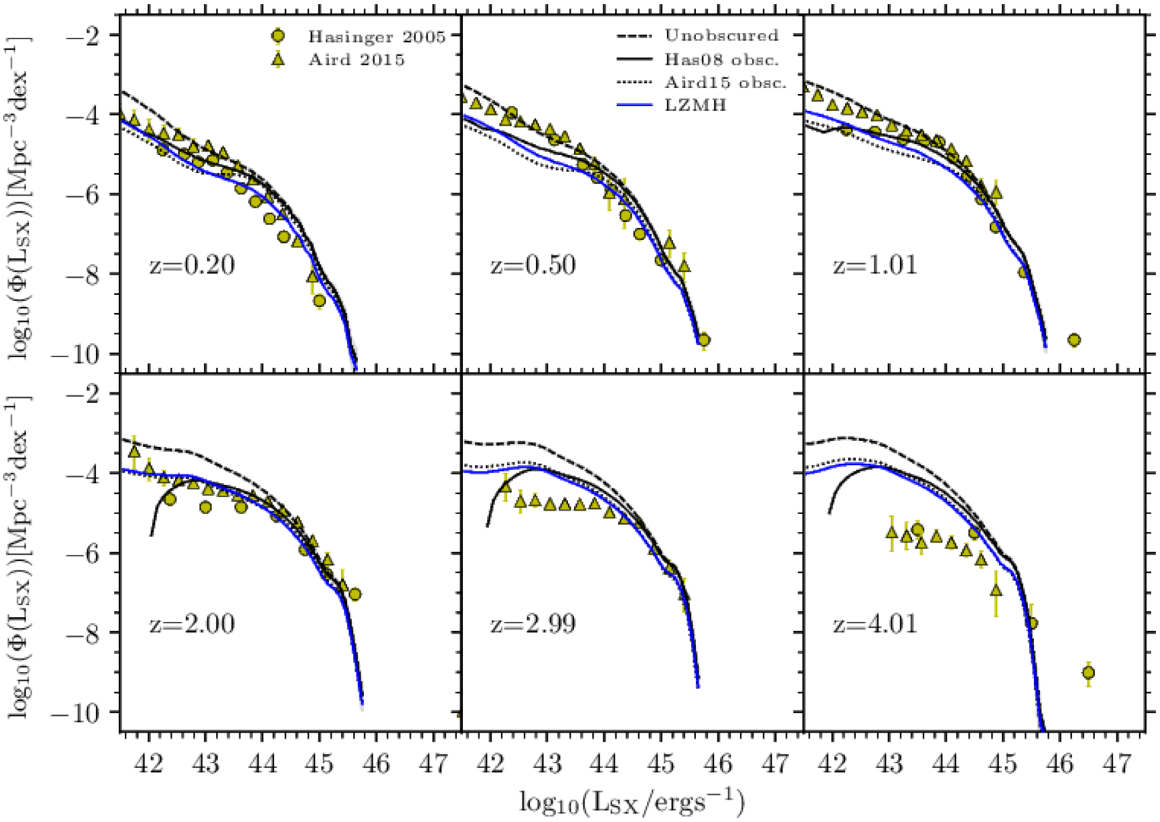

Figure 18. The predicted rest-frame soft X-ray luminosity function compared to observations. The dashed black line shows the prediction without accounting for absorption effects, the solid black line is the prediction using the Hasinger (2008) visible fraction, the dotted black line is using the Aird et al. (2015) visible fraction and the blue line is using our observationally determined LZMH visible fraction. The observations are Hasinger, Miyaji & Schmidt (2005) (circles) and Aird et al. (2015) (triangles).

bolometric luminosity function, whereas the model of Hirschmann et al. (2014) underpredicts the observed bolometric luminosity function by 1 dex at this redshift and luminosity. The hard X-ray luminosity function predicted by EAGLE in Rosas-Guevara et al. (2016) is compared to the observational estimate of Aird et al. (2015) over the redshift range $0 < z < 5$ in their fig. 7. Their model fits well to the observations at $z = 0$, but by $z = 1$, the slope of the luminosity function in their work is steeper than the observations. The model in this paper is in similar agreement for $z < 1$, and in better agreement with the observations for $z > 1$. For example, at $z = 2$, at $\log(L_{\rm HX}) = 10^{44}\,{\rm erg\,s^{-1}}$ (around the knee of the luminosity function at this redshift), our model agrees within 0.5 dex with the observations, whereas the model of Rosas-Guevara et al. (2016) underpredicts the observations by about 1 dex. Finally, Weinberger et al. (2018) compare the bolometric luminosity function from IllustrisTNG to Hopkins et al. (2007) in the redshift range $0 < z < 5$. Their model underpredicts the observations at the faint and bright end of the bolometric luminosity function and overpredicts the observations at intermediate luminosities at $z = 0.5$, and overpredicts the observations at all luminosities at $z = 3$. Around the knee of the luminosity function at $z = 3$ ($L_{\rm bol} = 3 \times 10^{46}\,{\rm ergs^{-1}}$), our model agrees within 0.5 dex with the observations, whereas the model of Weinberger et al. (2018) overpredicts the observations by 0.5 dex.

Overall, the AGN luminosity functions from the hydrodynamical simulations do not agree as well to the observational estimates as this model. The reasons for the differences in the AGN luminosity functions may be because the black hole mass accretion rates are calculated differently – in these simulations the Bondi-Hoyle approximation is used, as opposed to the calculation in Section 3.1 used in this work.

6 CONCLUSIONS

Understanding the evolution of AGNs across cosmic time has been of interest ever since they were discovered. AGNs have also been shown to be important in how galaxies evolve through AGN feedback. However, many uncertainties remain, such as the nature of the physical processes involved in AGN feedback.

We present predictions for the evolution of SMBHs and AGNs at $0 < z < 6$ using a high volume, high resolution dark matter simulation (P-Millennium) populated with galaxies using the semi-analytic model of galaxy formation GALFORM. This updated scheme for the SMBH spin evolution is used within the Lacey et al. (2016) GALFORM model as updated by Baugh et al. (2018) for the P-Millennium simulation. The Lacey et al. (2016) model has been shown to reproduce a large number of observable galaxy properties over an unprecedented wavelength and redshift range. The model that we use incorporates an updated prescription for SMBH spin evolution: for these predictions, we have assumed SMBH spin evolving in a 'chaotic accretion' scenario in which the angle between the accretion disc and the SMBH spin randomizes once a self-gravity mass of gas has been consumed.

We then calculated AGN bolometric luminosities from the SMBH mass accretion rate, taking into account the SMBH spin and the different radiative efficiencies for different accretion regimes (ADAFs,

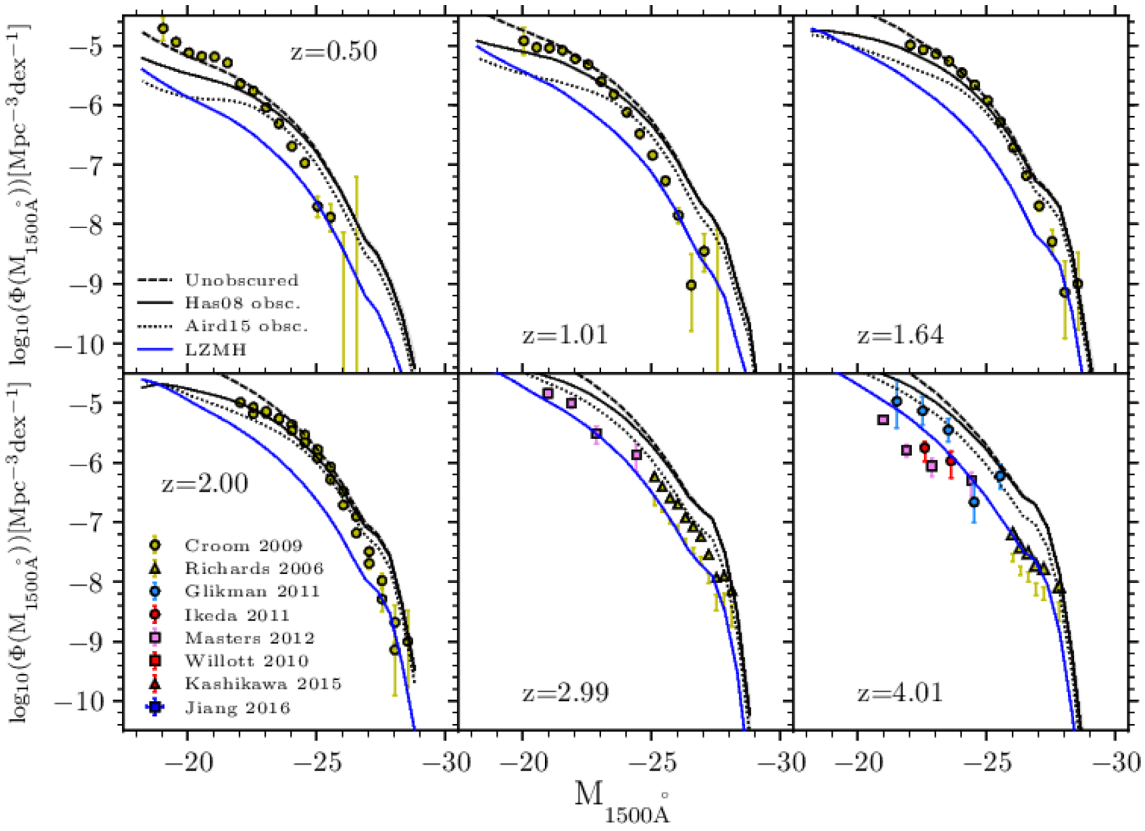

Figure 19. The predicted rest-frame 1500 Å luminosity function compared to observations that have been converted to 1500 Å. The dashed black line is the prediction without accounting for absorption effects, the solid black line is the prediction with the Hasinger (2008) visible fraction, the dotted black line is with the Aird et al. (2015) visible fraction and the blue line is with my observationally determined LZMH visible fraction. The observations are from SDSS DR3 – Richards et al. (2006) (yellow triangles); 2SLAQ + SDSS – Croom et al. (2009) (yellow circles); CFHQS + SDSS – Willott et al. (2010) (red squares); NDWFS + DLS – Glikman et al. (2011) (blue circles); the COSMOS field Ikeda et al. (2011) (red circles); Masters et al. (2012) (purple squares); Subaru – Kashikawa et al. (2015) (red triangles); and SDSS Stripe 82 – Jiang et al. (2016) (blue squares).

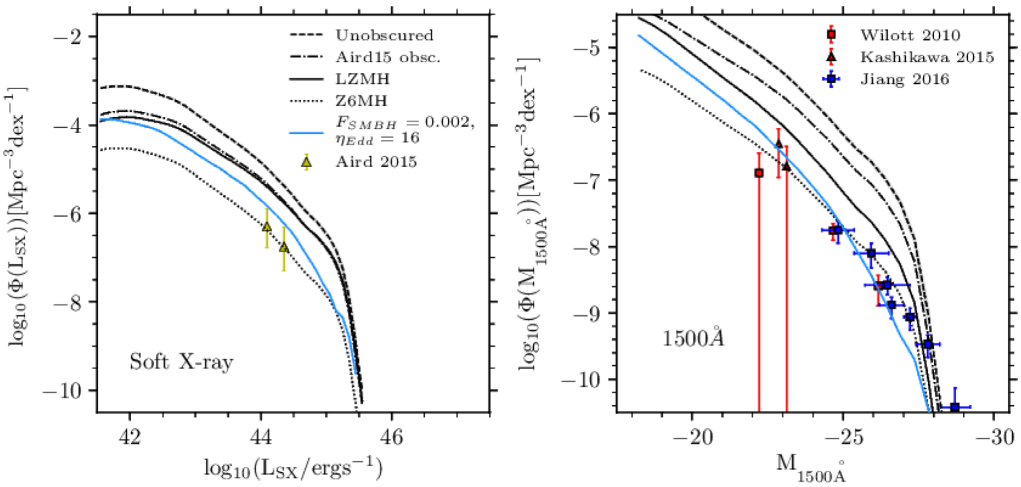

Figure 20. The rest-frame soft X-ray luminosity function (left-hand panel) and the rest-frame 1500 Å luminosity function (right-hand panel), both at $z = 6$. We show predictions without obscuration (dashed black), with the Aird et al. (2015) visible fraction (dot–dash), with the 'low-z modified Hopkins' (LZMH) visible fraction with the standard model (black solid), with the '$z = 6$ modified Hopkins' (Z6MH) visible fraction (black dotted) and with the 'low-z modified Hopkins' visible fraction with the different parameters (blue solid). The observations for the soft X-ray band are from Aird et al. (2015) (yellow triangles), and for 1500 Å are from Willott et al. (2010) (red squares), Kashikawa et al. (2015) (red triangles), and Jiang et al. (2016) (blue squares).

thin discs, super Eddington objects). Then using a template SED and different obscuration models we derived AGN luminosities in the hard X-ray, soft X-ray, and optical/UV (1500 Å) bands.

The model predictions are consistent with both the observed black hole mass functions and SMBH mass versus bulge mass correlations. We present the spin distribution of SMBHs in the simulation, for the chaotic and prolonged modes of accretion, and their evolution for $0 < z < 6$. The median SMBH spin in both the chaotic and prolonged modes evolves very little. For the prolonged mode, the scatter in the SMBH spin distribution decreases with increasing redshift. We also present examples of the evolution of spin and mass for typical SMBHs, and find that for most masses, the evolution is similar, except at the highest masses, $M_{\rm BH} > 10^8\,{\rm M}_\odot$, where mergers cause the SMBHs to be spun up to higher spin values.

We compare the AGN luminosity functions in the redshift range $0 < z < 6$ to a wide range of observations at different wavelengths. The model is in good agreement with the observations. We split the luminosity functions by accretion mode (ADAFs, thin discs, super-Eddington objects) and by fuelling mode (hot halo or starbursts triggered by disc instabilities or mergers) to see the relative contributions. At low redshifts, $z < 2$, and low luminosities, $L_{\rm bol} < 10^{43}\,{\rm erg\,s^{-1}}$, the ADAF contribution dominates but at higher luminosities and higher redshifts, the thin disc and super-Eddington objects dominate the luminosity function. Hot halo mode fuelled accretion dominates at $z < 3$, and $L_{\rm bol} < 10^{44}\,{\rm ergs^{-1}}$, but at higher redshift and higher luminosity, starbursts triggered by disc instabilities dominate the luminosity function.

There are many natural continuations from this work. We have already mentioned that we always assume a quasar SED for our bolometric corrections, while in reality we have a variety of AGNs having different accretion rates in different accretion regimes, which will have different SED shapes (e.g. Jin, Ward & Done 2012). Using different template SEDs for different regimes may allow the model to predict luminosity functions in better agreement with the observations. Secondly, we could more thoroughly explore the dependence of the model on the SMBH spin evolution model used e.g. investigating the dependence of the results on the size of the increments in mass used in the SMBH accretion calculation. Finally, in this paper, we do not show radio luminosity functions – given that AGN jets are observed to have a strong effect on their host galaxies and given that these jets emit at radio wavelengths via synchrotron emission, an investigation into radio emission would also be important for understanding the role of AGNs in galaxy evolution.

ACKNOWLEDGEMENTS

This work was supported by the Science and Technology facilities Council (STFC) ST/L00075X/1. AJG acknowledges an STFC studentship funded by grant ST/N50404X/1. CL is funded by an Australian Research Council Discovery Early Career Researcher Award (DE150100618) and by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. This work used the Distributed Research using Advanced Computing (DiRAC) Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC High Performace Computing Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grants ST/H008519/1 and ST/K00087X/1, STFC DiRAC Operations grant ST/K003267/1, and Durham University. DiRAC is part of the National E-Infrastructure.

REFERENCES

Abramowicz M. A., Czerny B., Lasota J. P., Szuszkiewicz E., 1988, ApJ, 332, 646
Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, MNRAS, 451, 1892
Barausse E., 2012, MNRAS, 423, 2533
Bardeen J. M., 1970, Nature, 226, 64
Bardeen J. M., Petterson J. A., 1975, ApJ, 195, L65
Bardeen J. M., Press W. H., Teukolsky S. A., 1972, ApJ, 178, 347
Baugh C. M. et al., 2018, MNRAS, 483, 4922
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Benson A. J., Borgani S., De Lucia G., Boylan-Kolchin M., Monaco P., 2012, MNRAS, 419, 3590
Berti E., Volonteri M., 2008, ApJ, 684, 822
Blandford R. D., Payne D. G., 1982, MNRAS, 199, 883
Blandford R. D., Znajek R. L., 1977, MNRAS, 179, 433
Blanton E. L., Randall S. W., Clarke T. E., Sarazin C. L., McNamara B. R., Douglass E. M., McDonald M., 2011, ApJ, 737, 99
Bonoli S., Marulli F., Springel V., White S. D. M., Branchini E., Moscardini L., 2009, MNRAS, 396, 423
Bournaud F., Combes F., Semelin B., 2005, MNRAS, 364, L18
Bournaud F., Elmegreen B. G., Elmegreen D. M., 2007, ApJ, 670, 237
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Boyle B. J., Fong R., Shanks T., Peterson B. A., 1990, MNRAS, 243, 1
Brenneman L. W., Reynolds C. S., 2006, ApJ, 652, 1028
Chiang C.-Y., Fabian A. C., 2011, MNRAS, 414, 2345
Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Collin-Souffrin S., Dumont A. M., 1990, A&A, 229, 292
Cowley W. I., Baugh C. M., Cole S., Frenk C. S., Lacey C. G., 2018, MNRAS, 474, 2352
Croom S. M. et al., 2009, MNRAS, 399, 1755
Croton D. J. et al., 2016, ApJS, 222, 22
De Lucia G., Fontanot F., Wilman D., Monaco P., 2011, MNRAS, 414, 1439
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Done C., Jin C., Middleton M., Ward M., 2013, MNRAS, 434, 1955
Dotti M., Colpi M., Pallini S., Perego A., Volonteri M., 2013, ApJ, 762, 68
Efstathiou G., Lake G., Negroponte J., 1982, MNRAS, 199, 1069
Elmegreen B. G., Bournaud F., Elmegreen D. M., 2008, ApJ, 688, 67
Enoki M., Ishiyama T., Kobayashi M. A. R., Nagashima M., 2014, ApJ, 794, 69
Fan X. et al., 2001, AJ, 122, 2833
Fanidakis N., Macciò A. V., Baugh C. M., Lacey C. G., Frenk C. S., 2013, MNRAS, 436, 315
Fanidakis N., Baugh C. M., Benson A. J., Bower R. G., Cole S., Done C., Frenk C. S., 2011, MNRAS, 410, 53
Fanidakis N. et al., 2012, MNRAS, 419, 2797
Fiacconi D., Sijacki D., Pringle J. E., 2018, MNRAS, 477, 3807
Forman W. et al., 2005, ApJ, 635, 894
Georgakakis A. et al., 2015, MNRAS, 453, 1946
Glikman E., Djorgovski S. G., Stern D., Dey A., Jannuzi B. T., Lee K.-S., 2011, ApJ, 728, L26
Gonzalez-Perez V. et al., 2018, MNRAS, 474, 4024
Hasinger G., 2008, A&A, 490, 905
Hasinger G., Miyaji T., Schmidt M., 2005, A&A, 441, 417
Hawley J. F., Balbus S. A., 2002, ApJ, 573, 738
Hawley J. F., Krolik J. H., 2006, ApJ, 641, 103
Hawley J. F., Gammie C. F., Balbus S. A., 1995, ApJ, 440, 742
Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A., Lacey C., 2003, MNRAS, 338, 903

Hirschmann M., Somerville R. S., Naab T., Burkert A., 2012, MNRAS, 426, 237
Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304
Hohl F., 1971, ApJ, 168, 343
Hopkins P. F., Richards G. T., Hernquist L., 2007, ApJ, 654, 731
Ikeda H. et al., 2011, ApJ, 728, L25
Jahnke K., Macciò A. V., 2011, ApJ, 734, 92
Jiang L. et al., 2009, AJ, 138, 305
Jiang L. et al., 2016, ApJ, 833, 222
Jiang L., Helly J. C., Cole S., Frenk C. S., 2014, MNRAS, 440, 2115
Jiang C. Y., Jing Y. P., Faltenbacher A., Lin W. P., Li C., 2008, ApJ, 675, 1095
Jin C., Ward M., Done C., 2012, MNRAS, 425, 907
Kashikawa N. et al., 2015, ApJ, 798, 28
Kauffmann G., Haehnelt M., 2000, MNRAS, 311, 576
Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188
Khandai N., Feng Y., DeGraf C., Di Matteo T., Croft R. A. C., 2012, MNRAS, 423, 2397
King A. R., Lubow S. H., Ogilvie G. I., Pringle J. E., 2005, MNRAS, 363, 49
King A. R., Pringle J. E., Hofmann J. A., 2008, MNRAS, 385, 1621
Kinney A. L., Schmitt H. R., Clarke C. J., Pringle J. E., Ulvestad J. S., Antonucci R. R. J., 2000, ApJ, 537, 152
Kudoh T., Matsumoto R., Shibata K., 1998, ApJ, 508, 186
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lacey C. G. et al., 2016, MNRAS, 462, 3854
Lagos C. D. P., Cora S. A., Padilla N. D., 2008, MNRAS, 388, 587
Lagos C. D. P., Padilla N. D., Cora S. A., 2009, MNRAS, 395, 625
Lagos C. d. P., Tobar R. J., Robotham A. S. G., Obreschkow D., Mitchell P. D., Power C., Elahi P. J., 2018, MNRAS, 481, 3573
Liu H., Wu Q., 2013, ApJ, 764, 17
Mahadevan R., 1997, ApJ, 477, 585
Malbon R. K., Baugh C. M., Frenk C. S., Lacey C. G., 2007, MNRAS, 382, 1394
Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, MNRAS, 351, 169
Marulli F., Bonoli S., Branchini E., Moscardini L., Springel V., 2008, MNRAS, 385, 1846
Masters D. et al., 2012, ApJ, 755, 169
McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184
McCullagh N., Norberg P., Cole S., Gonzalez-Perez V., Baugh C., Helly J., 2017, preprint(arXiv:1705.01988)
McKinney J. C., 2005, ApJ, 630, L5
Menci N., Fiore F., Lamastra A., 2013, ApJ, 766, 110
Menci N., Gatti M., Fiore F., Lamastra A., 2014, A&A, 569, A37
Mihos J. C., Hernquist L., 1996, ApJ, 464, 641
Natarajan P., Pringle J. E., 1998, ApJ, 506, L97
Neistein E., Netzer H., 2014, MNRAS, 437, 3373
Oke J. B., Gunn J. E., 1983, ApJ, 266, 713
Parkinson H., Cole S., Helly J., 2008, MNRAS, 383, 557
Peng C. Y., Impey C. D., Rix H.-W., Kochanek C. S., Keeton C. R., Falco E. E., Lehár J., McLeod B. A., 2006, ApJ, 649, 616
Penna R. F., Sądowski A., Kulkarni A. K., Narayan R., 2013, MNRAS, 428, 2255
Planck CollaborationXVI, 2014, A&A, 571, A16
Pounds K. A., Reeves J. N., King A. R., Page K. L., O'Brien P. T., Turner M. J. L., 2003, MNRAS, 345, 705
Pringle J. E., 1981, ARA&A, 19, 137
Randall S. W. et al., 2011, ApJ, 726, 86
Reeves J. N., O'Brien P. T., Ward M. J., 2003, ApJ, 593, L65
Rezzolla L., Barausse E., Dorband E. N., Pollney D., Reisswig C., Seiler J., Husa S., 2008, Phys. Rev. D, 78, 044002
Richards G. T. et al., 2006, AJ, 131, 2766
Rosas-Guevara Y., Bower R. G., Schaye J., McAlpine S., Dalla Vecchia C., Frenk C. S., Schaller M., Theuns T., 2016, MNRAS, 462, 190
Rupke D. S. N., Veilleux S., 2011, ApJ, 729, L27
Saha K., Cortesi A., 2018, ApJ, 862, L12
Sajina A., Yan L., Lacy M., Huynh M., 2007, ApJ, 667, L17
Schmidt M., 1968, ApJ, 151, 393
Schulze A., Wisotzki L., 2010, A&A, 516, A87
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, ApJ, 690, 20
Shankar F., Salucci P., Granato G. L., De Zotti G., Danese L., 2004, MNRAS, 354, 1020
Sharma P., Quataert E., Hammett G. W., Stone J. M., 2007, ApJ, 667, 714
Shirakata H. et al., 2019, MNRAS, 482, 4846
Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575
Simha V., Cole S., 2017, MNRAS, 472, 1392
Soltan A., 1982, MNRAS, 200, 115
Springel V. et al., 2005, Nature, 435, 629
Sądowski A., Narayan R., Penna R., Zhu Y., 2013, MNRAS, 436, 3856
Tchekhovskoy A., McKinney J. C., 2012, MNRAS, 423, L55
Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886
Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, ApJ, 786, 104
Vestergaard M., Osmer P. S., 2009, ApJ, 699, 800
Volonteri M., 2010, A&AR, 18, 279
Volonteri M., Sikora M., Lasota J.-P., 2007, ApJ, 667, 704
Volonteri M., Dubois Y., Pichon C., Devriendt J., 2016, MNRAS, 460, 2979
Watarai K.-y., Fukue J., Takeuchi M., Mineshige S., 2000, PASJ, 52, 133
Weinberger R. et al., 2018, MNRAS, 479, 4056
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Willott C. J. et al., 2010, AJ, 139, 906
Wu X.-B. et al., 2015, Nature, 518, 512
Younger J. D., Hopkins P. F., Cox T. J., Hernquist L., 2008, ApJ, 686, 815
Yuan F., Narayan R., 2014, ARA&A, 52, 529
Yuan F., Quataert E., Narayan R., 2003, ApJ, 598, 301

APPENDIX A: EFFECTS OF VARYING SMBH SEED MASS

In Fig. A1, we show the effect of varying the SMBH seed mass on the black hole mass function at $z = 0$. We show plots for SMBH seed masses of $10\,h^{-1}\,\mathrm{M}_\odot$ (the default value), $10^3\,h^{-1}\,\mathrm{M}_\odot$, and $10^5\,h^{-1}\,\mathrm{M}_\odot$. Generally, the black hole mass function reaches a converged value at about 100 times the black hole seed mass. We also plot as vertical lines $M_{\mathrm{BH}} = M_{\mathrm{seed}}$, $M_{\mathrm{BH}} = 2 \times M_{\mathrm{seed}}$, and $M_{\mathrm{BH}} = 3 \times M_{\mathrm{seed}}$. It can be seen that the spikes in the black hole mass function occur at these values due to SMBH seeds merging with other SMBH seeds.

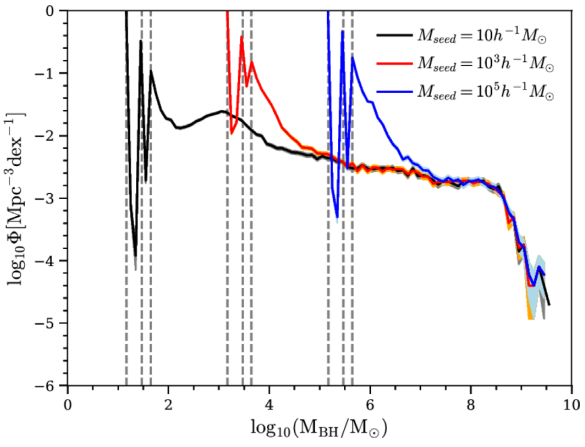

Figure A1. The black hole mass function at $z = 0$ for seed masses of $10h^{-1}\,\mathrm{M}_\odot$ (black), $10^3\,h^{-1}\,\mathrm{M}_\odot$ (red), and $10^5\,h^{-1}\,\mathrm{M}_\odot$ (blue).

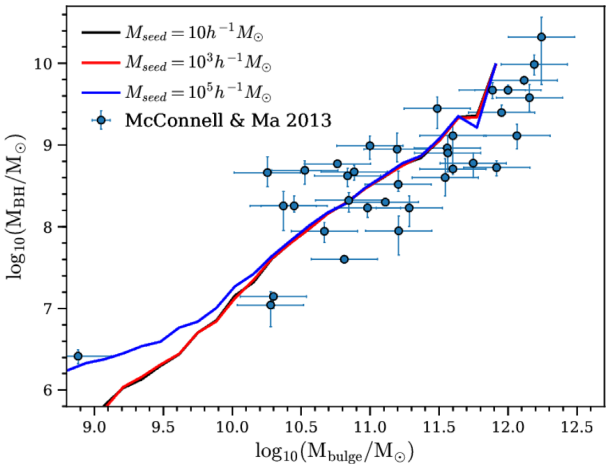

Figure A2. The predicted SMBH mass versus SMBH mass relation at $z = 0$ for seed masses of $10\,h^{-1}\,M_\odot$ (black), $10^3\,h^{-1}\,M_\odot$ (red), and $10^5\,h^{-1}\,M_\odot$ (blue) compared to McConnell & Ma (2013).

This convergence in properties at around 100 times the seed mass can also be seen in Fig. A2, where the median of the SMBH mass versus bulge mass relation for seeds of mass $10^5\,h^{-1}\,M_\odot$ only converges with that for the other seed masses for SMBH masses above around $10^7\,M_\odot$.

APPENDIX B: CALCULATING BROAD-BAND OPTICAL MAGNITUDES FOR AGNS

We define the filter-averaged luminosity per unit frequency for a filter R in the observer frame at redshift z as

$$< L_\nu >_R^{(z)} = \frac{\int L_\nu((1+z)\nu_o)R(\nu_o)\mathrm{d}\nu_o}{\int R(\nu_o)\mathrm{d}\nu_o}, \quad \text{(B1)}$$

where $L_\nu(\nu)$ is the luminosity per unit frequency in the rest frame and $R(\nu_o)$ is the response function of the filter at observed frequency ν_o. The absolute magnitude in the AB system in the observer frame band defined by the filter R for redshift z is then defined as

$$M_{AB,R}^{(z)} = -2.5\log_{10}\left(\frac{< L_\nu >_R^{(z)}}{L_{\nu_o}}\right), \quad \text{(B2)}$$

where $L_{\nu_o} = 4\pi(10\mathrm{pc}^2) \times f_{\nu_o}$ with $f_{\nu_o} = 3631\mathrm{Jy}$, the flux corresponding to an apparent AB magnitude of 0, and L_{ν_o} the corresponding luminosity per unit frequency for an absolute AB magnitude of 0. We remind readers that monochromatic AB (Absolute Bolometric) apparent magnitudes are defined using the following relation (Oke & Gunn 1983):

$$m_{\mathrm{AB}}(\nu) = -2.5\log_{10}\left(\frac{f_\nu}{f_{\nu_o}}\right), \quad \text{(B3)}$$

where f_ν is the observed flux of the source, which is related to the luminosity per unit frequency in the rest frame of the object as

$$f_\nu(\nu_o) = \frac{(1+z)L_\nu((1+z)\nu_o)}{4\pi d_L^2}. \quad \text{(B4)}$$

The apparent and observer frame absolute magnitudes for a filter R are then related by

$$m_{AB}(\nu) = -2.5\log_{10}\left(\frac{< L_\nu >_R^{(z)}}{L_{\nu_o}}\right) - 2.5\log_{10}(1+z) + 5\log_{10}(d_\mathrm{L}/10\,\mathrm{pc}). \quad \text{(B5)}$$

We then use the following formulae to convert the observational data from the different wavelengths given to rest-frame wavelength 1500 Å. Note that we are only comparing continuum luminosities in this study, which is consistent with the Marconi et al. (2004) template used throughout this paper. The data presented in the studies that we use have the contribution from the emission lines removed and so this is an appropriate comparison. The results presented in Richards et al. (2006) are given in the K-corrected SDSS i band at $z = 2$, which we write as $M_i'(z = 2)$. This is given by $M_i'(z = 2) = M_i(z = 2) - 2.5\log(1 + z)$, where we define $M_i(z = 2)$ as the absolute magnitude at the rest-frame wavelength corresponding to the observed i-band at $z = 2$, as in equations (B2) and (B5). To convert from $M_i(z = 2)$ to 1500 Å, we follow Richards et al. (2006) by using $L_\nu \propto \nu^{\alpha_\nu}$ but using a spectral index value of $\alpha_\nu = -0.44$ from Marconi et al. (2004) instead of $\alpha_\nu = -0.5$ in Richards et al. (2006). First, we convert from $M_i'(z = 2)$ to $M_i(z = 0)$ using equations (B1) and (B2):

$$\begin{aligned} M_i(z=0) &= M_i'(z=2) + 2.5(1+\alpha_\nu)\log(1+2) \\ &= M_i'(z=2) + 0.668, \end{aligned} \quad \text{(B6)}$$

where $M_i(z = 0)$ is the absolute magnitude at the central wavelength of the rest-frame i band (7471 Å) corresponding to equation (B2) for $z = 0$. Then we relate $M_i(z = 0)$ to the absolute magnitude at rest-frame 1500 Å, M_{1500}, to give the conversion to $M_i'(z = 2)$:

$$\begin{aligned} M_{1500} &= M_i(z=0) + 2.5\alpha_\nu \log_{10}\left(\frac{1500\,\text{Å}}{7471\,\text{Å}}\right), \\ &= M_i(z=0) + 0.767, \\ &= M_i'(z=2) + 1.435. \end{aligned} \quad \text{(B7)}$$

where in the last line we used equation (B6). Jiang et al. (2009), Willott et al. (2010), Ikeda et al. (2011), Masters et al. (2012), and Kashikawa et al. (2015) report observed absolute continuum magnitudes, M_{1450}, corresponding to rest frame 1450 Å. These absolute magnitudes are defined without the extra redshift factor included in the Richards et al. (2006) definition. These absolute magnitudes at 1450 Å, M_{1450}, can be converted to 1500 Å using

$$\begin{aligned} M_{1500} &= M_{1450} + 2.5\alpha_\nu \log_{10}\left(\frac{1500\,\text{Å}}{1450\,\text{Å}}\right), \\ &= M_{1450} - 0.016. \end{aligned} \quad \text{(B8)}$$

Finally, Croom et al. (2009) report observations in the SDSS g band (4670 Å) K-corrected to $z = 2$, so we use the correction in their paper:

$$M_g'(z=2) = M_i'(z=2) + 2.5\alpha_\nu \log\left(\frac{4670\,\text{Å}}{7471\,\text{Å}}\right), \quad \text{(B9)}$$

and combine it with the above relation to give

$$M_{1500} = M_g'(z=2) + 1.211. \quad \text{(B10)}$$

APPENDIX C: VISIBLE AND OBSCURED FRACTIONS FOR AGNS

The AGN visible fractions (the fraction of sources at a particular luminosity and redshift that are unobscured) derived in this paper have been estimated by constructing an observational bolometric luminosity function from observed luminosity functions at X-ray and optical wavelengths. These luminosities were converted to

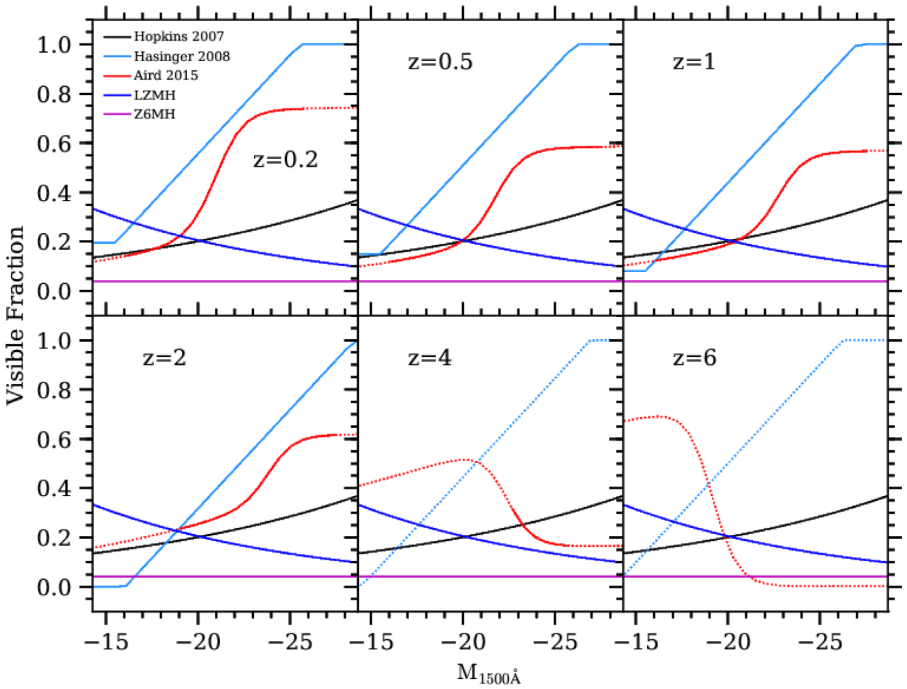

Figure C1. Comparing the visible fractions for rest-frame 1500 Å for different obscuration models. Shown are Hopkins et al. (2007) (black), Hasinger (2008) (light blue), Aird et al. (2015) (red), the LZMH model (dark blue), and the Z6MH model (purple). The solid lines for the observational visible fractions indicate the ranges where there is observational data, while the dotted lines indicate ranges where a functional form has been extrapolated.

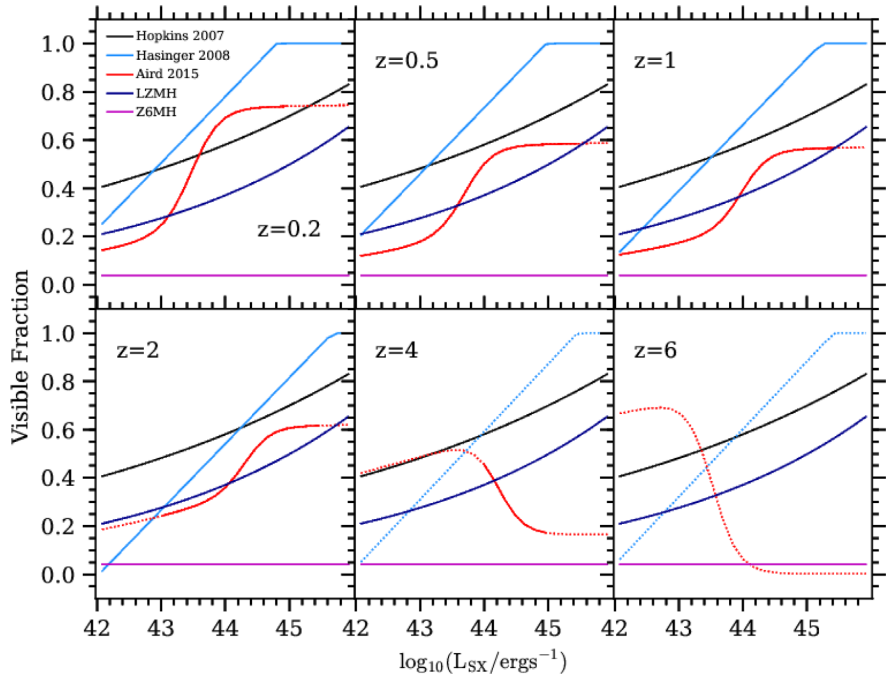

Figure C2. The same as the previous plot, but for rest-frame soft X-rays.

bolometric using the Marconi et al. (2004) AGN SED, and then the observed number densities were converted to total number densities using visible fractions of a functional form similar to Hopkins et al. (2007) dependent only on $L_{\rm bol}$ (cf. equation 40). We assumed that there is no obscuration for hard X-ray wavelengths. The coefficients in the expressions for the visible fractions were then selected (cf. equations 43–45) so as to minimize the scatter in the estimated bolometric luminosity function.

To construct a bolometric luminosity function from multiple sets of observations in different wavebands, different authors use different template SEDs. Some authors include reprocessed radiation from dust (its inclusion causes an 'IR bump' in the

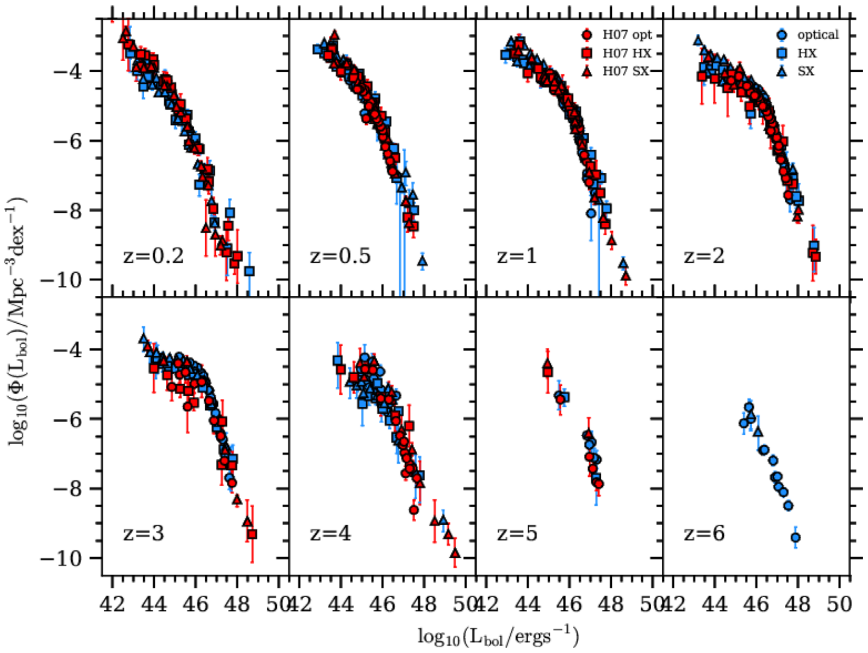

Figure C3. The bolometric luminosity function derived in this work (blue) by using the Marconi et al. (2004) bolometric corrections, and by varying the coefficients of the visible fractions to obtain a bolometric luminosity function with the smallest scatter between points derived from data at different wavelengths, compared to the Hopkins et al. (2007) bolometric luminosity function (red). The Hopkins et al. (2007) bolometric luminosities have been multiplied by 7.9/11.8 to account for the different SED template used (see the text).

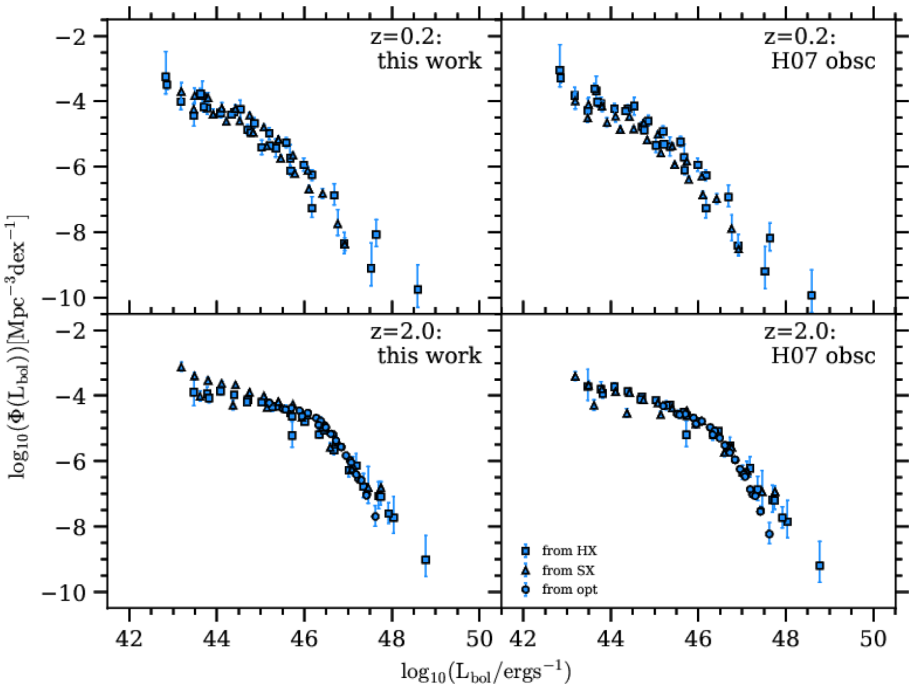

Figure C4. Comparing the effect of using different obscuration models on the constructed bolometric luminosity functions. The left-hand panels are obtained using the obscuration model presented in this work, while the right-hand panels use the obscuration model of Hopkins et al. (2007). The upper panels are for $z = 0.2$ and the lower panels are for $z = 2$.

SED) whereas some do not. Including reprocessed radiation gives observed bolometric luminosities, whereas not including the IR bump gives intrinsic bolometric luminosities. The intrinsic bolometric luminosities are isotropic, while the observed bolometric luminosities are not isotropic because the obscuring torus is not isotropic. The observed bolometric luminosity functions of Hopkins et al. (2007) are given in terms of observed rather than intrinsic bolometric luminosities, so when we compare with these, we multiply the luminosities of Hopkins et al. (2007) by a factor of 7.9/11.8 (cf. Marconi et al. 2004) to account for this effect.

We show a comparison of the different obscuration models at 1500 Å in Fig. C1 and at soft X-ray energies in Fig. C2. The values from different studies are not all on a single curve, and so there is clearly still some uncertainty in the visible fraction.

Our bolometric luminosity function is shown compared to the bolometric luminosity functions estimated in Hopkins et al. (2007) in Fig. C3, and the two are in agreement. The bolometric luminosity function derived in this work is also similar to that determined by Shankar et al. (2009).

Our observationally estimated visible fractions are redshift independent by construction. We have explored whether a better fit could be obtained by including a redshift dependence. To obtain a better fit, the visible fraction needs to increase and then decrease with redshift (cf. the redshift dependence derived by Aird et al. 2015), but even with a functional form to allow this, the scatter in the bolometric luminosity function was only slightly less than for redshift independent versions of the visible fraction.

To quantify the effect of using the new visible fraction derived in this paper, we compare the bolometric luminosity function derived using the Hopkins et al. (2007) visible fraction, to the bolometric luminosity function derived using the visible fraction presented in this paper, in Fig. C4. The new visible fraction does improve the constructed bolometric luminosity function, this reduction in scatter can be seen particularly at $L_{\rm bol} \sim 10^{44}$ erg s^{-1} at $z = 0.2$ and at $L_{\rm bol} \sim 10^{48}$ erg s^{-1} at $z = 2$.

APPENDIX D: THE EFFECT OF THE TIME AVERAGING METHOD

In this appendix, we show the effect of varying $\Delta t_{\rm window}$ on the AGN luminosity function, as introduced in Section 3.4, and compare the luminosity function obtained using the time averaging method in Section 3.4 to a luminosity function constructed using the snapshot luminosities. In Fig. D1, the predicted luminosity function with a value of $\Delta t_{\rm window} = t_{\rm H}/10$ (the standard model) is compared to the predicted luminosity function with a value of $\Delta t_{\rm window} = t_{\rm H}/50$. The two are very similar, except at low luminosities at high redshift, where there is a slight difference. The similarity shows that the value of $\Delta t_{\rm window}$ adopted does not strongly affect the luminosity function. In Fig. D2, the predicted luminosity function with a value

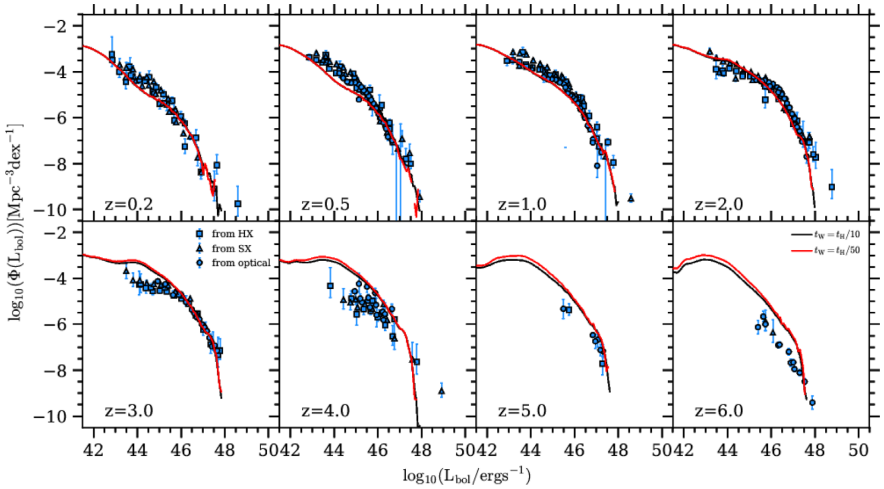

Figure D1. Exploring the effect on the AGN bolometric luminosity function of varying $\Delta t_{\rm window}$, shown are $\Delta t_{\rm window} = t_{\rm H}/10$ (black) and $\Delta t_{\rm window} = t_{\rm H}/50$ (red).

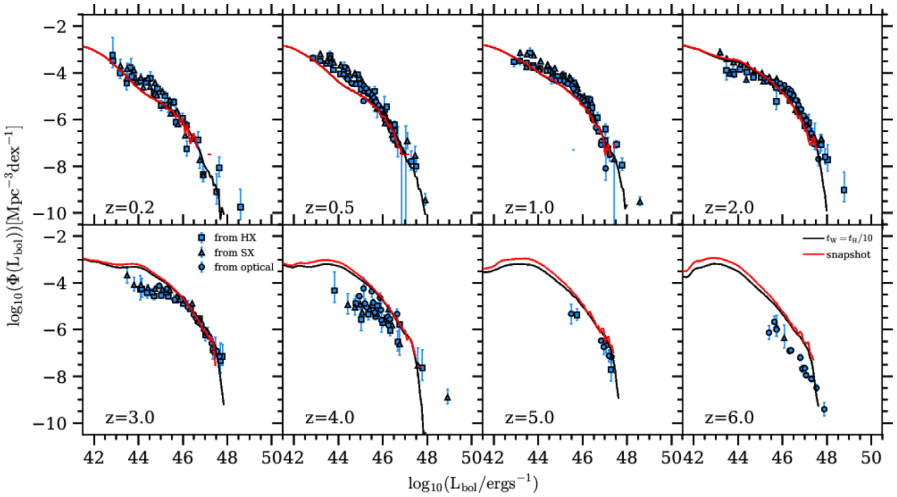

Figure D2. Exploring the effect on the AGN bolometric luminosity function of varying $\Delta t_{\rm window}$, shown are $\Delta t_{\rm window} = t_{\rm H}/10$ (black) and using the snapshot luminosities (red).

of $\Delta t_{\rm window} = t_{\rm H}/10$ is compared to the luminosity function where only the snapshot luminosities are used to construct the luminosity function. It can be seen how the time averaging method allows predictions for much lower number densities that for the snapshot case. These two cases are very similar in the luminosity range where they overlap, showing that the time averaging method does not change the predicted luminosity function significantly.

APPENDIX E: EXPLORING THE EFFECT OF VARYING PARAMETERS

We show the effect on the bolometric luminosity function of varying some of the free parameters for SMBH and AGN used in the model; in Fig. E1, we show the effect of varying the parameter $f_{\rm q}$ (cf. equation 26). $f_{\rm q}$ affects the value of $\dot{M}$ and therefore the AGN luminosities. One expects a higher value of $f_{\rm q}$ to lead to lower values of $\dot{M}$ and therefore a steeper luminosity function at the bright end, as we see in Fig. E1. At the faint end, a lower value of $f_{\rm q}$ results in a poorer fit to the observations at low redshift ($z = 0.2$, 0.5, 1) but is a better fit to the observations at high redshift ($z = 2$, 4, 6). At the bright end, a higher value of $f_{\rm q}$ seems to give a better fit to the observations at low redshift but gives a worse fit to the observations at high redshift (e.g. around $L_{\rm bol} \sim 10^{48}\,{\rm erg\,s^{-1}}$ at $z = 4$). With these considerations in mind, we decide to keep the Fanidakis et al. (2012) value of $f_{\rm q} = 10$ for our predictions in this paper.

We show the effect of varying the parameter $\eta_{\rm Edd}$ (cf. equation 31) in Fig. E2. $\eta_{\rm Edd}$ controls the suppression of the luminosity for super-Eddington accretion rates, where a low value of $\eta_{\rm Edd}$ corresponds to stronger luminosity suppression than a high value of $\eta_{\rm Edd}$. This parameter only affects the very bright end of the luminosity function, as we would expect. This parameter also has more of an effect at high redshift, where there are more super-Eddington sources. A value of $\eta_{\rm Edd} = 1$ gives a slightly better fit to the bright end observations at $z = 6$ but $\eta_{\rm Edd} = 16$ gives a better fit to bright end observations at $z = 2$ and $z = 4$. Therefore, we once again to opt to keep the Fanidakis et al. (2012) value of $\eta_{\rm Edd} = 4$ for our predictions in this paper.

We show the effect of switching off disc instabilities in Fig. E3. We show the fiducial model alongside a model in which all discs are stable and so no disc instability starbursts occur. Disc instabilities dominate the AGN luminosity function at $z > 2$, and so this is the regime where we expect turning off disc instabilities to have the most effect. For $L_{\rm bol} < 10^{46}\,{\rm erg\,s^{-1}}$, at $z > 2$ switching off disc instabilities results in fewer starbursts and so there are fewer objects at these luminosities. For $L_{\rm bol} > 10^{46}\,{\rm erg\,s^{-1}}$, at $z > 2$ the two models are similar – this is because if we switch off disc instabilities, galaxy mergers trigger the starbursts that would have otherwise happened due to disc instabilities. At $z < 2$, switching off disc instabilities makes the luminosity function less steep.

We show the effect of switching off the accretion and merger spin-up in Fig. E4. The radiative accretion efficiency given to the black holes is $\epsilon = 0.1$. The luminosity functions for the two models are generally similar, although the fiducial model has a slightly lower number density at high luminosities.

We show the effect of changing the assumptions for accretion efficiency, ϵ, in Fig. E5. We compare the fiducial model to a model in which the accretion efficiency is the thin disc accretion efficiency for all values of the specific mass accretion rate, $\dot{m}$. Interestingly, this result provides a slightly better fit to the bolometric luminosity function, particularly for $z < 0.5$ and $L_{\rm bol} < 10^{45}\,{\rm erg\,s^{-1}}$, where the fiducial model underpredicts the number density. This is the regime where ADAFs dominate the luminosity function, and so this test suggests that a better fit to the observed AGN luminosity function might be obtained if the radiative accretion efficiency for ADAFs is higher than the values assumed in our standard model.

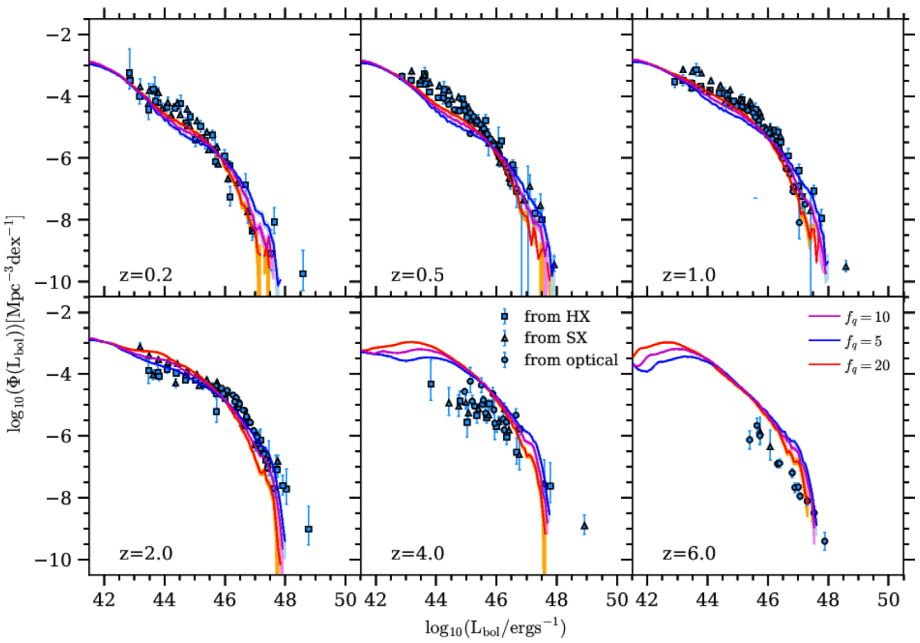

Figure E1. Exploring the effect on the AGN bolometric luminosity function of varying the parameter $f_{\rm q}$. Shown are $f_{\rm q} = 5$ (blue), $f_{\rm q} = 10$ (purple, the fiducial model), and $f_{\rm q} = 20$ (red). The shading shows the Poisson errors of the distribution.

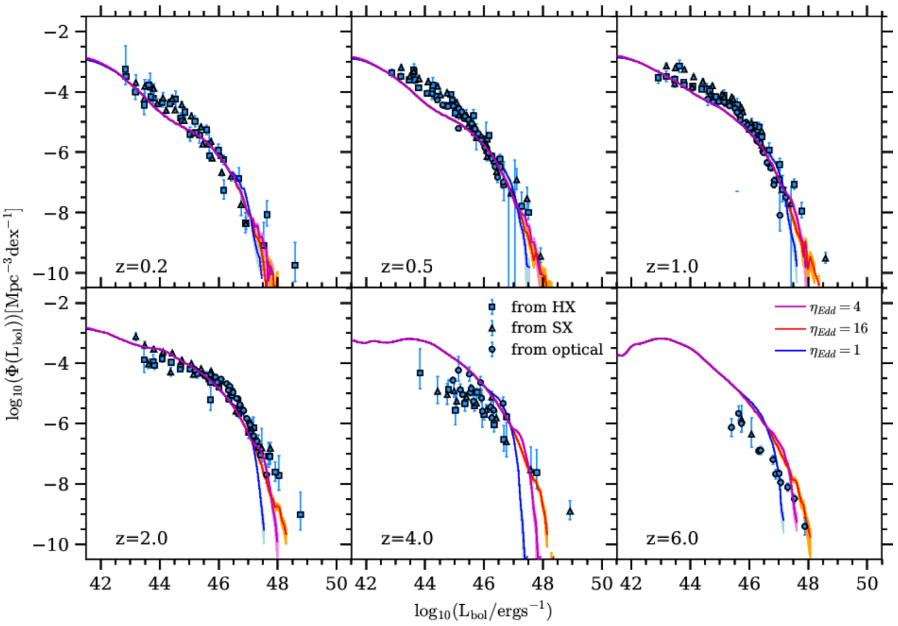

Figure E2. Exploring the effect of varying $\eta_{\rm Edd}$. Shown are $\eta_{\rm Edd} = 1$ (blue), $\eta_{\rm Edd} = 4$ (purple, the fiducial model), and $\eta_{\rm Edd} = 16$ (red).

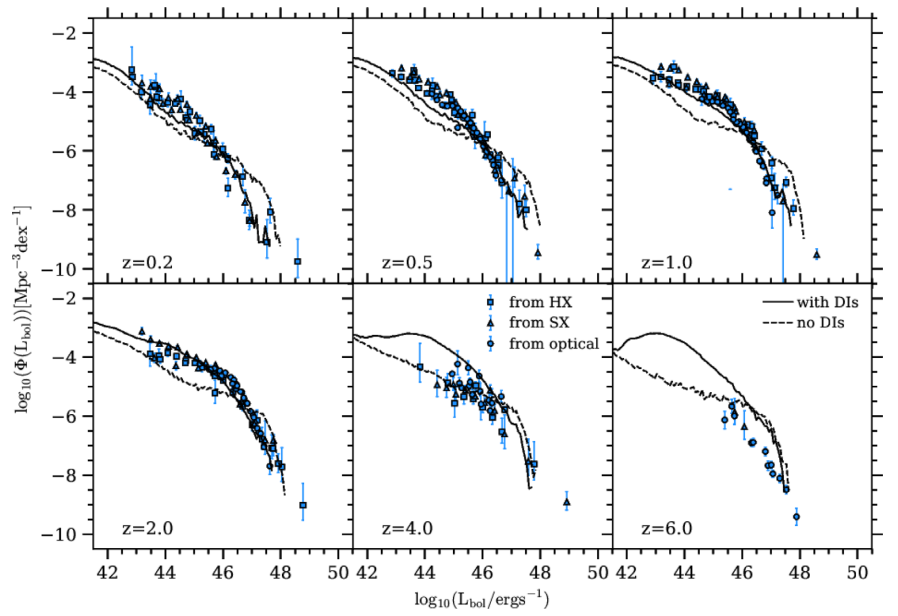

Figure E3. Exploring the effect of switching off disc instabilities. Shown are the fiducial model (solid) and the model with disc instabilities switched off (dashed).

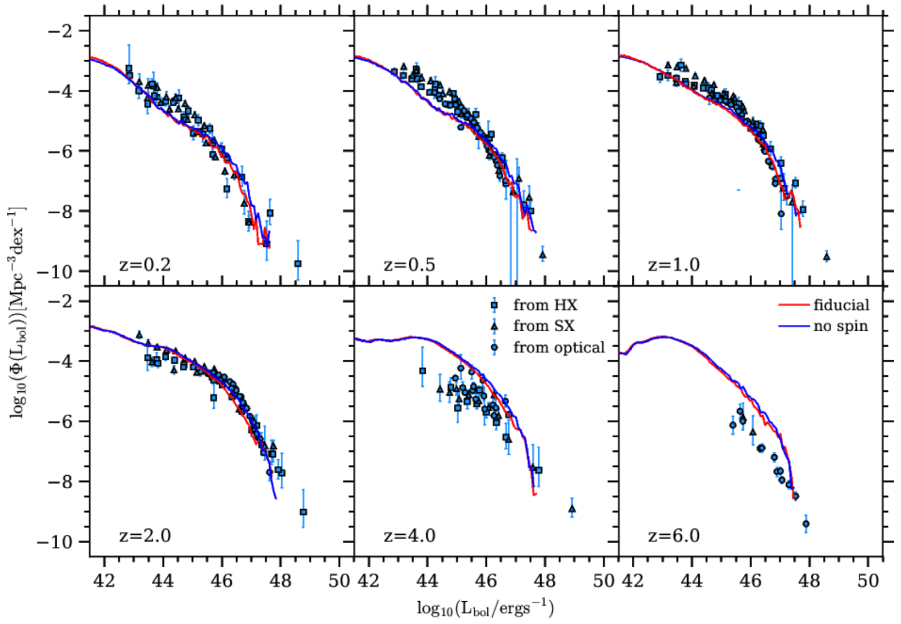

Figure E4. Exploring the effect of turning off the SMBH spin-up evolution: the model with chaotic mode accretion spin-up and merger spin-up (red) and the model with no accretion nor merger spin-up with a thin disc accretion efficiency, $\epsilon_{\rm TD} = 0.1$ (blue).

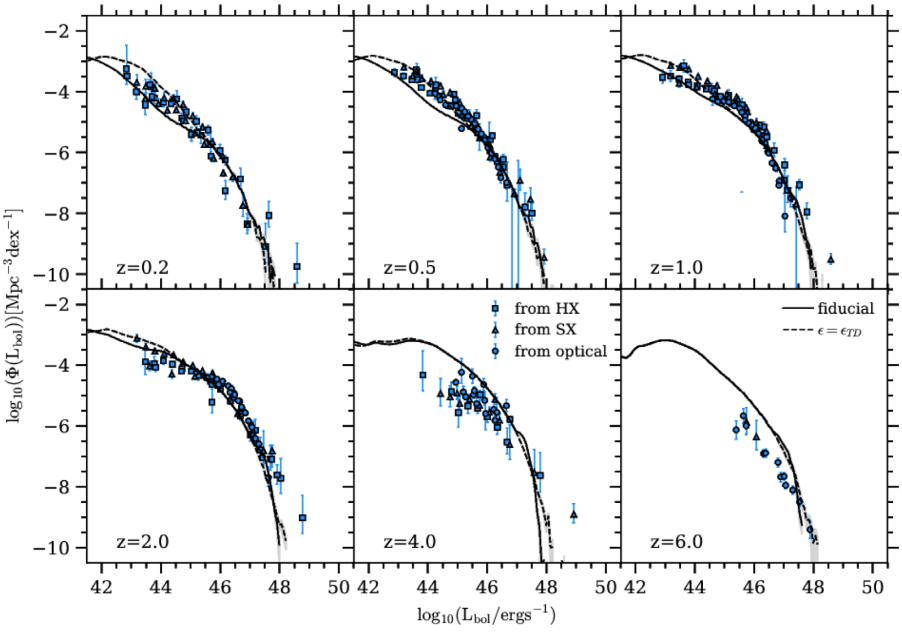

Figure E5. Exploring the effect of changing the accretion efficiency ϵ: the model with $\epsilon = \epsilon_{\rm TD}$ as the accretion efficiency for all $\dot{m}$ regimes (black dashed) and the fiducial model (black solid).

This paper has been typeset from a TEX/LATEX file prepared by the author.